

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 19:46:15 ; Search time 7.72727 Seconds
(without alignments)
205.411 Million cell updates/sec

Title: US-09-546-136-1.
Perfect score: 52
Sequence: 1 MKQAFVPEFD 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues 1107863

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	10	AAW37236	Partial sequence o
2	34	65.4	374	AAB30325	Thermus filiformis
3	34	65.4	406	AAB94069	Human protein sequ
4	34	65.4	537	AAB95734	Human protein sequ
5	34	65.4	1332	AAV01084	Human IKAP protein
6	34	65.4	1332	AAB93160	Human protein sequ
7	34	65.4	1332	AAB66345	Human IKAP. Homo
8	34	65.4	1332	ABB98492	Wild-type human IK
9	34	65.4	1332	ABB98493	Mutant human IKBA

10	34	65.4	1332	23	AAO15505	Human I-kappa beta
11	33	63.5	99	24	ABP78981	N. gonorrhoeae ami
12	33	63.5	534	22	ABB61497	Drosophila melanog
13	33	63.5	536	21	AAB44555	Virulence gene pro
14	33	63.5	536	22	ABP34507	Pasteurella multic
15	33	63.5	657	22	ABG14461	Novel human diagno
16	32	61.5	13	22	AAM98932	Vaccine related MH
17	32	61.5	36	21	AAG61119	Arabidopsis thalia
18	32	61.5	214	20	AAV37603	Chlamydia trachoma
19	32	61.5	214	23	ABG91055	Chlamydia trachoma
20	32	61.5	552	23	ABB91137	Herbicidally activ
21	32	61.5	789	22	ABB59220	Drosophila melanog
22	32	61.5	789	22	ABB60269	Drosophila melanog
23	32	61.5	1268	22	ABG29502	Novel human diagno
24	32	61.5	1372	22	ABG24404	Novel human diagno
25	32	61.5	1744	22	ABG62890	Drosophila melanog
26	31	59.6	45	22	ABG55298	Human liver peptid
27	31	59.6	45	22	ABB40125	Peptide #7631 enco
28	31	59.6	45	22	ABB24592	Protein #6591 enco
29	31	59.6	45	22	AAM60886	Human brain expres
30	31	59.6	45	22	AAM73569	Human bone marrow
31	31	59.6	45	22	AAM20021	Peptide #6455 enco
32	31	59.6	45	22	AAM33764	Peptide #7801 enco
33	31	59.6	45	23	ABG43435	Human peptide enco
34	31	59.6	109	18	AAW47068	Human brain P2X-1
35	31	59.6	218	23	ABP40187	Staphylococcus epi
36	31	59.6	228	20	AAV34572	Porphyromonas ging
37	31	59.6	229	20	AAV34571	Porphyromonas ging
38	31	59.6	231	20	AAV34429	Porphyromonas ging
39	31	59.6	329	21	AAG13171	Arabidopsis thalia
40	31	59.6	369	21	AAG41236	Arabidopsis thalia
41	31	59.6	388	18	AAW47066	Human brain P2X-1
42	31	59.6	388	19	AAW55035	HPRR amino acid s
43	31	59.6	388	22	AAB84382	Amino acid sequenc
44	31	59.6	399	17	AAW04218	Human urinary blad
45	31	59.6	404	18	AAW47067	Human brain P2X-2

ALIGNMENTS

RESULT 1
AAW37236

ID AAW37236 standard; peptide; 10 AA.

XX AAW37236;

AC AAW37236;

DT 17-JUN-1998 (first entry)

XX Partial sequence of botulinum neurotoxin serotype E associated protein.

DE Clostridia: botulinum toxin; neurotoxin; serotype E; vaccine;
KW antibody; acetylcholine; presynaptic nerve.

XX Clostridium botulinum.

OS WO9801754-Al.

PN 15-JAN-1998.

XX 08-JUL-1996; 96WO-US11383.

XX 08-JUL-1996; 96WO-US11383.

XX (UYMA-) UNIV MASSACHUSETTS DARTMOUTH.

XX Singh BR, Zhang Z;

XX WPI; 1998-101196/09.

XX Polypeptide complex synthesised by Clostridia - useful in, e.g.
PT producing vaccines for protecting animals against serotype E
PT neurotoxin

XX Claim 17; Page 3; 38pp; English.

PS A new peptide complex synthesised by Clostridia bacteria comprises

CC the serotype E botulinum neurotoxin and five neurotoxin associated

CC peptides of molecular weights 118, 80, 65, 40 and 18 kDa respectively.

CC The present sequence represents a partial sequence of the 18 kDa peptide.

CC Also claimed are: (1) detecting the serotype E neurotoxin complex in

CC biological samples by: (a) contacting with an antibody or polypeptide

CC specifically binding to a polypeptide of the complex, and (b) detecting

CC antibody- or polypeptide-bound polypeptide, indicating the presence of

CC serotype E neurotoxin; (2) an antibody binding to the complex, or one

CC of the five neurotoxin associated polypeptides, and (3) polypeptides

CC isolated from the complex. The complex can be used to produce

CC therapeutics for treating diseases resulting from excessive

CC acetylcholine release from presynaptic nerve terminals causing

CC undesirable contraction of smooth or skeletal muscle cells and resulting

CC in, e.g. spasmodic torticollis, essential tremor, spasmodic dysphonia,

CC charley horse, strabismus, blepharospasm, oromandibular dystonia,

CC spasms of the sphincters of the cardiovascular, gastrointestinal or

CC urinary systems, tardive dyskinesia, profuse sweating, lacrimation or

CC mucous secretion. The complex can also be used to produce therapeutics

CC for treating spasticity occurring secondary to brain ischaemia,

CC traumatic injury of the brain or spinal cord, tension headaches, pain

CC caused by sporting injuries or arthritic contractions. The complex may

CC further be used, within a carrier to prepare vaccines, useful for, e.g.

CC vaccinating animals against serotype E neurotoxin. The polypeptides can

CC be used for testing foodstuff, or diagnosing vertebrate

CC gastrointestinal, blood or tissue samples, for infection with serotype E

CC neurotoxin, useful in preventing botulism.

XX Sequence 10 AA;

Query Match 100.0%; Score 52; DB 19; Length 10;

Best Local Similarity 100.0%; Pred. NO. 0.00037;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKQAFVFEPD 10

DB 1 |||||

RESULT 2

AAB30325

ID AAB30325 standard; Protein; 374 AA.

XX AAB30325;

XX 12-FEB-2001 (first entry)

XX Thermus filiformis Tfil methylase.

XX Tfil: methylase; restriction endonuclease; nucleic acid cleavage.

XX Thermus filiformis.

XX US6133008-A.

XX 17-OCT-2000.

XX 07-MAY-1999; 99US-0306881.

XX 07-MAY-1999; 99US-0306881.

XX (NEW) NEW ENGLAND BIOLABS INC.

XX Xu S, Hsieh P;

XX WPI; 2001-006139/01.

XX N-PSDB; AAC65519.

XX New DNA encoding a Thermus filiformis restriction endonuclease for

XX creating or producing recombinant molecules and for producing large

PT quantities of this enzyme

XX Disclosure: Column 11-14; 14pp; English.

XX The present invention describes a methylase and restriction endonuclease

CC from Thermus filiformis. These are involved in the bacterial

CC restriction-modification system. The genes and protein sequences can be

CC used in the laboratory for the cleavage of nucleic acids.

XX Sequence 374 AA;

Query Match 65.4%; Score 34; DB 22; Length 374;

Best Local Similarity 50.0%; Pred. NO. 87;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKQAFVFEPD 10

DB 326 MKKGFIYTFD 335

RESULT 3

AAB94069

ID AAB94069 standard; Protein; 406 AA.

XX AAB94069;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:14254.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs.

XX Claim 8; SEQ ID 14254; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesising 5602

XX full-length cDNAs defined in the specification. Where a primer set

XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides, and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in

XX the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 406 AA;

Query Match 65.4%; Score 34; DB 22; Length 406;
 Best Local Similarity 60.0%; Pred. No. 95;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKQAFVFEFD 10
 : | | | | |
 Db 314 LKVLFLFEFD 323

RESULT 4
 AAB95734
 ID AAB95734 standard; Protein; 537 AA.

XX AC AAB95734;

XX DT 26-JUN-2001 (first entry)

XX DE Human protein sequence SEQ ID NO:18619.

XX KW Human; primer: detection; diagnosis; antisense therapy; gene therapy.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI: 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602
 full-length cDNAs defined in the specification, and for the detection
 and/or diagnosis of the abnormality of the proteins encoded by the
 full-length cDNAs.

XX PS Claim 8; SEQ ID 18619; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesising 5602
 full-length cDNAs defined in the specification. Where a primer set
 comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 to the complementary strand of a polynucleotide which comprises one of
 the 5602 nucleotide sequences defined in the specification, where the
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 of an oligonucleotide comprising a sequence complementary to the
 complementary strand of a polynucleotide which comprises a 5'-end
 sequence and an oligonucleotide comprising a sequence complementary to a
 polynucleotide which comprises a 3'-end sequence, where the
 oligonucleotide comprises at least 15 nucleotides and the combination of
 the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense thera
 CC in gene therapy. The primers are useful for synthesising polynucle
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 537 AA;

Query Match 65.4%; Score 34; DB 22; Length 537;
 Best Local Similarity 60.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKQAFVFEFD 10
 : | | | | |
 Db 445 LKVLFLFEFD 454

RESULT 5
 AAY01084
 ID AAY01084 standard; Protein; 1332 AA.

XX AC AAY01084;

XX DT 09-JUN-1999 (first entry)

XX DE Human IKAP protein sequence.

XX KW IKAP; human; cellular signal transduction regulator;

XX KW transcriptional activation regulator; translatable transcript.

XX OS Homo sapiens.

XX PN US5891719-A.

XX PD 06-APR-1999.

XX PF 16-NOV-1997; 97US-0971244.

XX PR 16-NOV-1997; 97US-0971244.

XX PA (TULA-) TULARIK INC.

XX PI Bacuerle P, Cohen L;

XX DR WPI: 1999-253865/21.

XX DR N-PSDB: AAX28025.

XX PT Isolated, recombinant cDNA sequences encoding human IKAP
 polypeptides - useful for regulating cellular signal transduction
 and transcriptional activation

XX PS Claim 12; Column 9-18; 15pp; English.

XX This sequence is the human IKAP protein of the invention. IKAP
 regulates cellular signal transduction and transcriptional activation
 (e.g. the activation of nuclear factor kappaB). Nucleic acids encoding
 CC IKAP have a range of applications including use as translatable
 CC transcripts, hybridisation probes, polymerase chain reaction (PCR)
 CC primers, diagnostic nucleic acids, use in detecting the presence of IKAP
 CC genes and gene transcripts and in detecting or amplifying nucleic acids
 CC encoding additional IKAP homologues and structural analogues. In
 CC diagnosis IKAP hybridisation probes may be used to identify wild-type
 CC and mutant IKAP alleles in clinical samples. In therapy, IKAP nucleic
 CC acids can be used to modulate cellular expression or intracellular
 CC concentration or availability of active IKAP.

XX Sequence 1332 AA;

As a consequence

sequence 1332 AA;

KW Human; IKBKAP; Familial Dysautonomia; FD; Riley-Day syndrome;
 KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening.
 OS Homo sapiens.
 PN WO200259381-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 07-JAN-2002; 2002WO-US00473.
 XX
 PR 06-JAN-2001; 2001US-260080P.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Staugenhaupt S, Gusella JF;
 XX
 DR WPI: 2002-674806/72.
 DR N-PSDB; ABQ80569.
 XX
 PT New IKBKAP genes with mutations, useful for identifying a subject with
 PT familial dysautonomia (FD), or for rapid carrier screening in the
 PT Ashkenazi Jewish population, e.g. screening presymptomatic homozygotes
 PT or prenatal diagnosis
 XX
 PS Claim 8; Fig 8; 109pp; English.
 XX
 CC The present invention relates to methods and compositions useful for
 CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
 CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) [OMIM
 CC 223900]. It was found that mutations in the IKBKAP gene (see ABQ80565)
 CC are associated with FD. The mutation associated with the major haplotype
 CC of FD, FDI mutation, is a base pair (bp) mutation, where the thymine
 CC nucleotide located at bp 6 of intron 20 in the IKBKAP gene is replaced
 CC with a cytosine. This results in skipping of exon 20 in the mRNA from FD
 CC patients, although they continue to express varying levels of wild-type
 CC message in a tissue-specific manner. The mutation associated with the
 CC minor haplotype, FD2 mutation, is a bp mutation, where the guanine
 CC nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a cytosine.
 CC This bp mutation causes an arginine to proline missense mutation (R696P)
 CC in the IKBKAP protein, which is predicted to disrupt a potential
 CC phosphorylation site. The IKBKAP nucleic acid sequences are useful for
 CC identifying a subject with FD and for rapid carrier screening. The IKBKAP
 CC gene maps to chromosome 9q31. The present sequence is the wild-type human
 CC IKBKAP protein sequence.
 XX
 SQ Sequence 1332 AA;
 Query Match 65.4%; Score 34; DB 23; Length 1332;
 Best Local Similarity 60.0%; Pred. No. 3.4e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MKQAFVFEFD 10
 : | | : | | | | |
 Db 1240 LKVLFLFEFD 1249
 RESULT 9
 ABB98493
 ID ABB98493 standard; Protein: 1332 AA.
 XX
 AC ABB98493;
 XX
 DT 08-NOV-2002 (first entry)
 XX
 DE Mutant human IKBKAP.
 XX
 XX Human; IKBKAP; Familial Dysautonomia; FD; Riley-Day syndrome; mutant;
 KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;
 KW mtein.
 XX
 OS Homo sapiens.
 XX

PH Key Location/Qualifiers
 FT Misc-difference 696 /note= "Wild-type Arg replaced by Pro"
 XX
 PN WO200259381-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 07-JAN-2002; 2002WO-US00473.
 XX
 PR 06-JAN-2001; 2001US-260080P.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Staugenhaupt S, Gusella JF;
 XX
 DR WPI: 2002-674806/72.
 DR N-PSDB; ABQ80570.
 XX
 PT New IKBKAP genes with mutations, useful for identifying a subject with
 PT familial dysautonomia (FD), or for rapid carrier screening in the
 PT Ashkenazi Jewish population, e.g. screening presymptomatic homozygotes
 PT or prenatal diagnosis
 XX
 PS Claim 9; Page -: 109pp; English.
 XX
 CC The present invention relates to methods and compositions useful for
 CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
 CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) [OMIM
 CC 223900]. It was found that mutations in the IKBKAP gene (see ABQ80565)
 CC are associated with FD. The mutation associated with the major haplotype
 CC of FD, FDI mutation, is a base pair (bp) mutation, where the thymine
 CC nucleotide located at bp 6 of intron 20 in the IKBKAP gene is replaced
 CC with a cytosine. This results in skipping of exon 20 in the mRNA from FD
 CC patients, although they continue to express varying levels of wild-type
 CC message in a tissue-specific manner. The mutation associated with the
 CC minor haplotype, FD2 mutation, is a bp mutation, where the guanine
 CC nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a cytosine.
 CC This bp mutation causes an arginine to proline missense mutation (R696P)
 CC in the IKBKAP protein, which is predicted to disrupt a potential
 CC phosphorylation site. The IKBKAP nucleic acid sequences are useful for
 CC identifying a subject with FD and for rapid carrier screening. The IKBKAP
 CC gene maps to chromosome 9q31. The present sequence is the mutant human
 CC IKBKAP protein sequence.
 CC Note: the present sequence was not shown in the specification, but was
 CC derived from the human wild-type IKBKAP sequence given in Fig 8.
 XX
 SQ Sequence 1332 AA;
 Query Match 65.4%; Score 34; DB 23; Length 1332;
 Best Local Similarity 60.0%; Pred. No. 3.4e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MKQAFVFEFD 10
 : | | : | | | | |
 Db 1240 LKVLFLFEFD 1249
 RESULT 10
 AAO15505
 ID AAO15505 standard; Protein: 1332 AA.
 XX
 AC AAO15505;
 XX
 DT 24-OCT-2002 (first entry)
 XX
 DE Human I-kappa beta associated protein (IKAP).
 XX
 XX Human; single nucleotide polymorphism; SNP; bronchial asthma;
 KW I-kappa beta associated protein; IKAP; infant bronchial asthma.
 XX
 OS Homo sapiens.
 XX

PN WO200259305-A1.
 PD 01-AUG-2002.
 XX 25-JAN-2002; 2002WO-JP00540.
 XX 25-JAN-2001; 2001JP-0017076.
 PR (SAKA) OTSUKA PHARM CO LTD.
 XX Nakamura Y, Tamari M;
 XX WPI; 2002-557950/59.
 DR N-PSDB; AAL44190.
 XX
 XX Detection of specific single nucleotide polymorphisms in human I-kappa
 PT beta associated protein for analysis of etiology of bronchial asthma -
 XX
 XX Claim 3; Page 39-45; 60pp; Japanese.
 XX
 XX The invention comprises a method for detecting single nucleotide
 CC polymorphisms (SNPs) in genes associated with human bronchial asthma. The
 CC method specifically refers to detecting polymorphisms in the gene
 CC encoding human I-kappa beta associated protein (IKAP). The invention also
 CC comprises primers and probes for use in the method of the invention. The
 CC method of the invention is useful for analysis of the etiology of
 CC bronchial asthma (especially infant bronchial asthma). The present amino
 CC acid sequence represents the human I-kappa beta associated protein.
 XX
 XX Sequence 1332 AA;
 SQ
 Query Match 65.4%; Score 34; DB 23; Length 1332;
 Best Local Similarity 60.0%; Pred. No. 3.4e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MKQAFVFEED 10
 DB 1240 LKVLFLFEED 1249
 RESULT 11
 ID ABP78981 standard; Protein; 99 AA.
 XX ABP78981;
 AC
 XX 07-MAR-2003 (first entry)
 DT
 XX N. gonorrhoeae amino acid sequence SEQ ID 4492.
 DE Antibacterial; infection; vaccine; gene therapy.
 XX Neisseria gonorrhoeae.
 OS
 XX WO200279243-A2.
 PN
 XX 10-OCT-2002.
 PD
 XX 12-FEB-2002; 2002WO-IB02069.
 PF
 XX 12-FEB-2001; 2001GB-0003424.
 PR (CHIR-) CHIRON SPA.
 XX Fontana MR, Pizza M, Massignani V, Monaci E;
 XX WPI; 2003-058415/05.
 DR N-PSDB; ABZ39951.
 XX
 XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 PT medicament for treating or preventing N. gonorrhoeae infection -
 XX
 PS Disclosure; Page 517; 815pp; English.

XX The present invention relates to proteins from Neisseria gonorrhoeae.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
 CC molecules of the invention.
 XX
 XX Sequence 99 AA;
 SQ
 Query Match 63.5%; Score 33; DB 24; Length 99;
 Best Local Similarity 60.0%; Pred. No. 33;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MKQAFVFEED 10
 DB 40 VORAFGEED 49
 RESULT 12
 ID ABB61497 standard; Protein; 534 AA.
 XX ABB61497;
 AC
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster polypeptide SEQ ID NO 11283.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR N-PSDB; ABL05600.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX Disclosure; SEQ ID NO 11283; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 534 AA;
 SQ
 Query Match 63.5%; Score 33; DB 22; Length 534;

```
Best Local Similarity 55.6%; Pred. No. 2.1e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 5; Conservative 2;

QY 1 MKQAFVFEF 9
Db 156 MRMGFIPEF 164

RESULT 13
AAB44555
ID AAB44555 standard; Protein; 536 AA.
XX
AC AAB44555;
XX
DT 08-FEB-2001 (first entry)
XX
DE Virulence gene protein #35.
XX
KW Virulence gene; antibacterial; vaccine; bacterial infection;
KW septicemia; bronchopneumonia; rhinitis; wound infection.
XX
OS Pasteurella multocida.
XX
PN WO200061724-A2.
XX
PD 19-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09218.
XX
PR 09-APR-1999; 99US-0128689.
PR 10-SEP-1999; 99US-0153453.
XX
PA (PHAA ) PHARMACIA & UPJOHN INC.
XX
PI Lowery DE, Fuller TE, Kennedy MJ;
DR WPI: 2000-647422/62.
DR N-PSDB; AAC79616.
XX
PT Attenuated Pasteurellaceae bacteria comprising mutations in virulence
genes, useful as a live attenuated vaccine against bacterial infections
PT
PS Claim 39; Pages 208-209; 322pp; English.
XX
CC The family Pasteurellaceae encompasses several pathogens that infect a
wide variety of animals. The present invention relates to virulence genes
from Pasteurellaceae. The virulence genes of the present invention may be
such virulence gene. The virulence genes of the present invention may be
mutated in order to produce an inactive gene. The inactive virulence gene
may in turn be used to produce a vaccine, which is useful for treating
bacterial infections such as septacemias, bronchopneumonias, rhinitis and
wound infections.
XX
SQ Sequence 536 AA;

Query Match 63.5%; Score 33; DB 21; Length 536;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 QAFVFEED 10
Db 234 QAILFEED 241

RESULT 14
ABP54507
ID ABP54507 standard; Protein; 536 AA.
XX
AC ABP54507;
XX
DT 24-JAN-2003 (first entry)
XX
```

```
DE Pasteurella multocoda yabk protein SEQ ID NO:71.
XX Antibacterial; vaccine; gram negative bacterial virulence gene;
KW identification; virulence; Pasteurellaceae.
XX
OS Pasteurella multocoda.
XX
PN WO200275507-A2.
XX
PD 26-SEP-2002.
XX
PF 17-JAN-2002; 2002WO-US01971.
XX
PR 15-MAR-2001; 2001US-0809665.
XX
PA (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI Lowery DE, Fuller TE, Kennedy MJ;
XX
DR WPI: 2002-740868/80.
DR N-PSDB; ABQ83493.
XX
PT New mutant gram-negative bacteria, useful as vaccines and for
identifying new anti-bacterial agents that target virulence genes and
their products
PT
XX Claim 36; Page 213-214; 350pp; English.
XX
CC The present invention describes a gram-negative bacteria comprising a
mutation in a gene, where the mutation results in decreased activity of
a gene product encoded by the mutated gene. Also described is a method
for producing a gram-negative bacteria mutant or an attenuated
Pasteurellaceae bacteria. The mutated genes have antibacterial activity
and can be used in vaccines. The gram-negative bacteria or the
attenuated Pasteurellaceae bacteria can be used as vaccines in the
fields of human medicine or veterinary medicine, and for identifying
new antibacterial agents that target the virulence genes and their
products. ABQ83458 to ABQ83578 and ABP54473 to ABP54551 represents
sequences used in the exemplification of the present invention.
XX
SQ Sequence 536 AA;

Query Match 63.5%; Score 33; DB 23; Length 536;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 QAFVFEED 10
Db 234 QAILFEED 241

RESULT 15
ABG14461
ID ABG14461 standard; Protein; 657 AA.
XX
AC ABG14461;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #14452.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
```

PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 XX WPI: 2001-639362/73.
 XX N-PSDB: AAS78648.
 DR
 DR
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID NO 44820; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 657 AA;

Query Match 63.5%; Score 33; DB 22; Length 657;
 Best Local Similarity 66.7%; Pred. No. 2.6e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKQAFVFEF 9
 |||||
 Db 132 MKSAFMFSF 140

Search completed: July 24, 2003, 20:05:57
 Job time : 9.72727 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:02:22 ; Search time 2.66234 Seconds
(without alignments)
158.924 Million cell updates/sec

Title: US-09-546-136-1

Perfect score: 52

Sequence: 1 MKQAFVFEFD 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	65.4	374	3	US-09-306-881-2
2	34	65.4	1332	2	US-08-971-244-2
3	34	65.4	1332	3	US-09-286-891-2
4	33	63.5	256	2	US-08-286-819A-41
5	33	63.5	256	3	US-08-280-337-41
6	31	59.6	218	4	US-09-134-001C-5032
7	31	59.6	237	4	US-09-107-532A-6194
8	31	59.6	388	2	US-08-742-621-1
9	31	59.6	388	3	US-09-191-608-22
10	31	59.6	399	2	US-08-742-621-3
11	31	59.6	399	2	US-08-750-134A-11
12	31	59.6	399	3	US-09-363-745-11
13	30	57.7	211	1	US-08-631-607-4
14	30	57.7	211	4	US-09-098-358B-4
15	30	57.7	232	3	US-08-906-769-81
16	30	57.7	232	3	US-08-906-616-81
17	30	57.7	232	3	US-08-817-795-81
18	30	57.7	232	3	US-08-639-075A-81
19	30	57.7	232	3	US-09-012-431-81
20	30	57.7	232	3	US-09-012-692-81
21	30	57.7	232	3	US-08-906-613-81
22	30	57.7	232	5	PCT-US95-14442A-81
23	30	57.7	261	2	US-08-484-905-111
24	30	57.7	261	3	US-08-481-985B-111
25	30	57.7	261	3	US-08-370-476-111
26	30	57.7	392	4	US-09-464-535-36
27	29	55.8	36	2	US-08-222-617A-18

28	29	55.8	160	4	US-09-198-452A-1244	Sequence 1,
29	29	55.8	204	1	US-08-155-171B-2	Sequence 2,
30	29	55.8	204	2	US-08-435-998-2	Sequence 2,
31	29	55.8	208	4	US-09-252-991A-32467	Sequence 32467,
32	29	55.8	268	4	US-09-107-532A-6057	Sequence 6057,
33	29	55.8	309	4	US-09-810-347-6	Sequence 6, Appl.,
34	29	55.8	311	4	US-09-328-352-6665	Sequence 6665, Ap
35	29	55.8	370	3	US-08-904-032-1	Sequence 1, Appli
36	29	55.8	399	2	US-08-742-621-4	Sequence 4, Appli
37	29	55.8	399	2	US-08-750-134A-5	Sequence 5, Appli
38	29	55.8	399	3	US-09-363-745-5	Sequence 5, Appli
39	29	55.8	903	4	US-09-339-159B-32	Sequence 32, Appl
40	28	53.8	96	4	US-09-387-418A-25	Sequence 25, Appl
41	28	53.8	114	4	US-09-328-352-5218	Sequence 5218, Ap
42	28	53.8	129	4	US-09-387-418A-24	Sequence 24, Appl
43	28	53.8	135	4	US-09-732-210-1115	Sequence 1115, Ap
44	28	53.8	172	4	US-08-737-109-17	Sequence 17, Appl
45	28	53.8	185	4	US-09-387-418A-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-09-306-881-2
; Sequence 2, Application US/09306881A
; Patent No. 6133008
; GENERAL INFORMATION:
; APPLICANT: HSIEN, Pei-chung
; APPLICANT: XU, Shuang-yong
; TITLE OF INVENTION: METHOD FOR CLONING AND PRODUCING THE TFIII RESTRICTION
; TITLE OF INVENTION: ENDONUCLEASE IN E. COLI
; FILE REFERENCE: NEB-159
; CURRENT APPLICATION NUMBER: US/09/306.881A
; CURRENT FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Thermus filiformis
US-09-306-881-2

Query Match 65.4%; Score 34; DB 3; Length 374;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 MKQAFVFEFD 10
Db 326 MKKGFYTFD 335

RESULT 2
US-08-971-244-2
; Sequence 2, Application US/08971244
; Patent No. 5891719
; GENERAL INFORMATION:
; APPLICANT: Cohen, Lucy
; APPLICANT: Baerle, Patrick
; TITLE OF INVENTION: IKAP Proteins, Nucleic Acids and Methods
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/971,244
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: OSMAN, RICHARD A
;; REGISTRATION NUMBER: 36,627
;; REFERENCE/DOCKET NUMBER: T97-011
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (650) 343-4341
;; TELEFAX: (650) 343-4342
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1332 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-971-244-2

Query Match 65.4%; Score 34; DB 2; Length 1332;

Best Local Similarity 60.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 MKQAFVFEFD 10
: | | :
Db 1240 LKVLFLFEFD 1249

RESULT 3
US-09-286-891-2
; Sequence 2, Application US/09286891
; Patent No. 6172195
; GENERAL INFORMATION:
; APPLICANT: Cohen, Lucy
; APPLICANT: Baeuerle, Patrick
; TITLE OF INVENTION: IKAP Proteins, Nucleic Acids and Methods
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/286,891
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/971,244
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1332 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-286-891-2

Query Match 65.4%; Score 34; DB 3; Length 1332;

Best Local Similarity 60.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Oy 1 MKQAFVFEFD 10
: | | :
Db 1240 LKVLFLFEFD 1249

RESULT 4
US-08-286-819A-41
; Sequence 41, Application US/08286819A
; Patent No. 5871910
; GENERAL INFORMATION:
; APPLICANT: ARTHUR, MICHEL
; APPLICANT: DUKTA-MALEN, SYLVIE
; APPLICANT: MOLINAS, CATHERINE
; APPLICANT: COURVALIN, PATRICE
; TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
; TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPOLYMERES, IN PARTICULAR
; TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
; TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,819A
; FILING DATE: 05-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/174,682
; FILING DATE: 28-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/917,146
; FILING DATE: 10-AUG-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR/91/00855
; FILING DATE: 29-OCT-1991
; APPLICATION DATA: FR 9013579
; FILING DATE: 31-OCT-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, No. 5871910man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-060-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELETYPE: 248655 OPAT UR
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 256 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; US-08-286-819A-41

Query Match

63.5%; Score 33; DB 2; Length 256;

Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KQAFVFEED 10
||| ||| |
Db 123 KQTFTEID 131

RESULT 5

US-08-980-357-41
; Sequence 41, Application US/08980357
; Patent No. 6013508

GENERAL INFORMATION:

; APPLICANT: ARTHUR, MICHEL

; APPLICANT: DUKTA-MALEN, SYLVIE

; APPLICANT: MOLINAS, CATHERINE

; APPLICANT: COURVALIN, PATRICE

; TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE

; TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPOLYMERES, IN PARTICULAR

; TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR

; TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS

; NUMBER OF SEQUENCES: 54

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,

; ADDRESS: P.C.

; STREET: 1755 S. Jefferson Davis Highway, Suite 400

; CITY: Arlington

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION NUMBER: US/08/980,357

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/286,819

; FILING DATE: 05-AUG-1994

; APPLICATION NUMBER: US 08/174,682

; FILING DATE: 28-DEC-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/917,146

; FILING DATE: 10-AUG-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/FR/91/00855

; FILING DATE: 29-OCT-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: FR 9013579f

; FILING DATE: 31-OCT-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Oblon, No. 6013508man F.

; REGISTRATION NUMBER: 24,618

; REFERENCE/DOCKET NUMBER: 660-060-0 PCT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 413-3000

; TELEFAX: (703) 413-2220

; INFORMATION FOR SEQ ID NO: 41:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 256 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

; ORIGINAL SOURCE:

; ORGANISM: Escherichia coli

; US-08-980-357-41

Query Match

63.5%; Score 33; DB 3; Length 256;

Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KQAFVFEED 10
||| ||| |
Db 123 KQTFTEID 131

RESULT 6

US-09-134-001C-5032
; Sequence 5032, Application US/09134001C
; Patent No. 6380370

GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 5032

; LENGTH: 218

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

; US-09-134-001C-5032

Query Match

59.6%; Score 31; DB 4; Length 218;

Best Local Similarity 60.0%; Pred. No. 80;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKQAFVFEED 10

||| ||| ||

Db 8 MYKAVVEFED 17

RESULT 7

US-09-107-532A-6194
; Sequence 6194, Application US/09107532A
; Patent No. 6583275

GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESSER: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6194:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc.feature
LOCATION: (B) LOCATION 1...237
SEQUENCE DESCRIPTION: SEQ ID NO: 6194:
US-09-107-532A-6194

Query Match 59.6%; Score 31; DB 4; Length 237;
Best Local Similarity 66.7%; Pred. No. 87;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKQAFVFEF 9
Db 143 MKMAFLPGF 151

RESULT 8
US-08-742-621-1
Sequence 1, Application US/08742621
Patent No. 5856129
GENERAL INFORMATION:
APPLICANT: HILLMAN, JENNIFER L.
APPLICANT: COLEMAN, ROGER
TITLE OF INVENTION: NOVEL HUMAN PURINOCEPTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,621
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0147 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY:
CLONE: consensus

Query Match 59.6%; Score 31; DB 2; Length 388;
US-08-742-621-1

Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 4 AFVFEFD 10
Db 10 AFLFEYD 16

RESULT 9
US-09-191-608-22
Sequence 22, Application US/09191608
Patent No. 6242216
GENERAL INFORMATION:
APPLICANT: Lynch, Kevin J.
APPLICANT: Burgard, Edward C.
APPLICANT: Metzger, Randy E.
APPLICANT: Niforatos, Wende
APPLICANT: Touma, Edward B.
APPLICANT: Van Biesen, T.
TITLE OF INVENTION: Nucleic Acids Encoding a Functional
TITLE OF INVENTION: Human Purinoceptor P2X2 and P2X4 And Methods Of Production
TITLE OF INVENTION: And Use Thereof
FILE REFERENCE: 6394 US.P1
CURRENT APPLICATION NUMBER: US/09/191,608
CURRENT FILING DATE: 1998-11-13
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 22
LENGTH: 388
TYPE: PRT
ORGANISM: Homo sapiens
US-09-191-608-22

Query Match 59.6%; Score 31; DB 3; Length 388;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AFVFEFD 10
Db 10 AFLFEYD 16

RESULT 10
US-08-742-621-3
Sequence 3, Application US/08742621
Patent No. 5856129
GENERAL INFORMATION:
APPLICANT: HILLMAN, JENNIFER L.
APPLICANT: COLEMAN, ROGER
TITLE OF INVENTION: NOVEL HUMAN PURINOCEPTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,621
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0147 US

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 399 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 166438
 US-08-742-621-3

Query Match 59.6%; Score 31; DB 2; Length 399;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AFVFEFD 10
 11:11:1
 Db 11 AFLFEYD 17

RESULT 11

US-08-750-134A-11
 Sequence 11, Application US/08750134A
 Patent No. 5985603
 GENERAL INFORMATION:
 APPLICANT: VALERA, SOLEDAD
 APPLICANT: BUELL, GARY
 TITLE OF INVENTION: P2X RECEPTORS (PURINOCEPTOR FAMILY)
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NIXON & VANDERHVE P.C.
 STREET: 1100 NORTH GLEBE ROAD
 CITY: ARLINGTON
 STATE: VIRGINIA
 COUNTRY: U.S.A.
 ZIP: 22201-4714
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA: US/08750.134A

APPLICATION NUMBER: US/08750.134A
 FILING DATE: 22-JAN-1997
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: CRAWFORD, ARTHUR C.
 REGISTRATION NUMBER: 25,327
 REFERENCE/DOCKET NUMBER: 1430-116
 TELEPHONE: (703) 816-4006
 TELEFAX: (703) 816-4100
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 399 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-750-134A-11

Query Match 59.6%; Score 31; DB 2; Length 399;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AFVFEFD 10
 11:11:1
 Db 11 AFLFEYD 17

RESULT 12

US-09-363-745-11
 Sequence 11, Application US/09363745
 Patent No. 6194162
 GENERAL INFORMATION:
 APPLICANT: VALERA, SOLEDAD
 APPLICANT: BUELL, GARY
 TITLE OF INVENTION: P2X RECEPTORS (PURINOCEPTOR FAMILY)
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NIXON & VANDERHVE P.C.
 STREET: 1100 NORTH GLEBE ROAD
 CITY: ARLINGTON
 STATE: VIRGINIA
 COUNTRY: U.S.A.
 ZIP: 22201-4714
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA: US/09/363.745

FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/750.134
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: CRAWFORD, ARTHUR C.
 REGISTRATION NUMBER: 25,327
 REFERENCE/DOCKET NUMBER: 1430-116
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4006
 TELEFAX: (703) 816-4100
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 399 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-363-745-11

Query Match 59.6%; Score 31; DB 3; Length 399;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AFVFEFD 10
 11:11:1
 Db 11 AFLFEYD 17

RESULT 13

US-08-631-607-4
 Sequence 4, Application US/08631607
 Patent No. 5767252
 GENERAL INFORMATION:
 APPLICANT: Worley et al, Paul
 TITLE OF INVENTION: NOVEL NEURONAL CELL GROWTH FACTOR
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 4225 Executive Square, Suite 1400
 CITY: La Jolla
 STATE: CA
 COUNTRY: USA
 ZIP: 92037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/631,607
FILING DATE: 08-APR-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/086001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 211 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-631-607-4

Query Match 57.7%; Score 30; DB 1; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KQAFVF 7
| | | | |
Db 6 KQAFVF 11

RESULT 14

US-09-098-358B-4
Sequence 4, Application US/09098358B
Patent No. 6436673
GENERAL INFORMATION:

APPLICANT: THE JOHN HOPKINS SCHOOL OF MEDICINE
APPLICANT: WORLEY, PAUL
APPLICANT: TSUI, CYNTHIA
TITLE OF INVENTION: NOVEL NEURONAL CELL GROWTH FACTOR
FILE REFERENCE: JHU 1340-1 (07265/086002)
CURRENT APPLICATION NUMBER: US/09/098,358B
CURRENT FILING DATE: 1998-06-16
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4

LENGTH: 211

TYPE: PRT

ORGANISM: ARTIFICIAL

FEATURE:

OTHER INFORMATION: RAT CRP

US-09-098-358B-4

Query Match 57.7%; Score 30; DB 4; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KQAFVF 7
| | | | |
Db 6 KQAFVF 11

RESULT 15

US-08-906-769-81
Sequence 81, Application US/08906769
Patent No. 6077687
GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
APPLICANT: Silver, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
MOLECULES AND USES THEREOF

NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/906,769
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA: 7

APPLICATION NUMBER: US 08/639,075

FILING DATE: 24-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER: 2618-25-C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 81:

SEQUENCE CHARACTERISTICS:

LENGTH: 232 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-906-769-81

Query Match 57.7%; Score 30; DB 3; Length 232;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKQAFVFEFD 10

: : | | | | |

Db 81 VEOTFVHQFD 90

Search completed: July 24, 2003, 20:09:23
Job time : 3.66234 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:08:50 ; Search time 4.28571 Seconds
(without alignments)
277.106 Million cell updates/sec

Title: US-09-546-136-1

Perfect score: 52

Sequence: 1 MKQAFVFEFD 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query % Match	Score	Length	ID	Description
1	36	69.2	749	15	US-10-099-352-40
2	34	65.4	1332	14	US-10-041-856-3
3	34	65.4	1332	14	US-10-041-856-5
4	33	63.5	538	11	US-09-820-843A-9
5	32	61.5	91	14	US-10-090-035-22
6	32	61.5	92	14	US-10-090-035-20
7	32	61.5	92	14	US-10-090-035-24
8	32	61.5	744	15	US-10-099-352-41
9	31	59.6	45	9	US-09-864-761-39890
10	31	59.6	388	10	US-09-833-082-2
11	31	59.6	431	15	US-10-156-761-9506
12	30	57.7	88	10	US-09-796-692-995
13	30	57.7	88	10	US-09-796-692-1524
14	30	57.7	88	15	US-10-040-862-995
15	30	57.7	88	15	US-10-040-862-1524

16	30	57.7	90	10	US-09-796-692-790	Sequence 790, App
17	30	57.7	90	15	US-10-040-862-790	Sequence 790, App
18	30	57.7	95	10	US-09-925-300-1214	Sequence 1214, Ap
19	30	57.7	97	10	US-09-796-692-1386	Sequence 1386, Ap
20	30	57.7	97	10	US-09-796-692-2111	Sequence 2111, Ap
21	30	57.7	97	15	US-10-040-862-1386	Sequence 1386, Ap
22	30	57.7	97	15	US-10-040-862-2111	Sequence 2111, Ap
23	30	57.7	181	15	US-10-156-761-9415	Sequence 9415, Ap
24	30	57.7	211	15	US-10-224-951-4	Sequence 4, Appli
25	30	57.7	297	9	US-09-925-302-794	Sequence 794, App
26	30	57.7	439	15	US-10-156-761-9626	Sequence 9626, Ap
27	30	57.7	1332	14	US-10-041-856-4	Sequence 4, Appli
28	30	57.7	2771	10	US-09-808-602-82	Sequence 82, Appl
29	30	57.7	2771	11	US-09-800-198-70	Sequence 70, Appl
30	29	56.7	433	15	US-10-128-714-8545	Sequence 8545, Ap
31	29	55.8	72	10	US-09-796-692-2362	Sequence 2362, Ap
32	29	55.8	72	15	US-10-040-862-2362	Sequence 2362, Ap
33	29	55.8	87	9	US-09-864-761-40699	Sequence 40699, A
34	29	55.8	182	9	US-09-864-761-36062	Sequence 36062, A
35	29	55.8	185	11	US-09-988-067B-56	Sequence 56, Appl
36	29	55.8	199	15	US-10-133-797-77	Sequence 77, Appl
37	29	55.8	246	9	US-09-815-242-13518	Sequence 13518, A
38	29	55.8	309	15	US-10-224-413-6	Sequence 6, Appli
39	29	55.8	331	9	US-09-803-286A-10	Sequence 10, Appl
40	29	55.8	436	15	US-10-190-279-5	Sequence 5, Appli
41	29	55.8	569	9	US-09-931-147-2	Sequence 2, Appli
42	29	55.8	569	15	US-10-133-797-2	Sequence 2, Appli
43	29	55.8	575	9	US-09-220-091-7	Sequence 7, Appli
44	29	55.8	646	9	US-09-841-132-567	Sequence 567, App
45	29	55.8	1247	9	US-09-803-286A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-10-099-352-40
; Sequence 40, Application US/10099352
; Publication No. US20030082569A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Clayton H.
; APPLICANT: York, J. Lyndal
; APPLICANT: McEwen, Joan E.
; TITLE OF INVENTION: Histoplasma Capsulation Catalase Sequences and Their Use in t
; TITLE OF INVENTION: of Histoplasma Capsulation and Histoplasmosis
; FILE REFERENCE: 40715-255988
; CURRENT APPLICATION NUMBER: US/10/099,352
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/275,353
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 749
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-099-352-40

Query Match 69.2%; Score 36; DB 15; Length 749;
Best Local Similarity 70.0%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 MKQAFVFEFD 10

DB 487 MKKAFSEFD 496

RESULT 2

US-10-041-856-3
; Sequence 3, Application US/10041856
; Publication No. US20020169299A1
; GENERAL INFORMATION:
; APPLICANT: SLAUGENIAUPT, SUSAN

; APPLICANT: GUSELLA, JAMES F.
; TITLE OF INVENTION: GENE FOR IDENTIFYING INDIVIDUALS WITH FAMILIAL
; FILE REFERENCE: 1829-4004US1
; CURRENT APPLICATION NUMBER: US/10/041,856
; PRIOR FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: 60/260,080
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-041-856-3

Query Match 65.4%; Score 34; DB 14; Length 1332;
Best Local Similarity 60.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKQAFVFEED 10
DB 1240 LKVLFLFEED 1249
:|:|:|

RESULT 3

US-10-041-856-5
; Sequence 5, Application US/10041856
; Publication No. US2002016929A1
; GENERAL INFORMATION:
; APPLICANT: SLAUGENHAUPT, SUSAN
; APPLICANT: GUSELLA, JAMES F.
; TITLE OF INVENTION: GENE FOR IDENTIFYING INDIVIDUALS WITH FAMILIAL
; FILE REFERENCE: 1829-4004US1
; CURRENT APPLICATION NUMBER: US/10/041,856
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: 60/260,080
; PRIOR FILING DATE: 2001-01-06
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-041-856-5

Query Match 65.4%; Score 34; DB 14; Length 1332;
Best Local Similarity 60.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKQAFVFEED 10
DB 1240 LKVLFLFEED 1249
:|:|:|

RESULT 4

US-09-820-843A-9
; Sequence 9, Application US/09820843A
; Publication No. US2003039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 538
; TYPE: PRT
; ORGANISM: H. influenzae

; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: thiamin ABC transporter, permease protein, putative
; NAME/KEY: misc_feature
; OTHER INFORMATION: gill574049
US-09-820-843A-9

Query Match 63.5%; Score 33; DB 11; Length 538;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 QAFVFEED 10
DB 234 QAILFEED 241
||:||||

RESULT 5

US-10-090-035-22
; Sequence 22, Application US/10090035
; Publication No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-090-035-22

Query Match 61.5%; Score 32; DB 14; Length 91;
Best Local Similarity 60.0%; Pred. No. 39;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKQAFVFEED 10
DB 32 VKETVVOEED 41
:|:|:|

RESULT 6

US-10-090-035-20
; Sequence 20, Application US/10090035
; Publication No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-090-035-20

Query Match 61.5%; Score 32; DB 14; Length 92;
Best Local Similarity 60.0%; Pred. No. 39;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKQAFVFEED 10
:|:|:|

Db 33 VKETVQEFED 42

RESULT 7

US-10-090-035-24

; Sequence 24, Application US/10090035

; Publication No. US20020170089A1

; GENERAL INFORMATION:

; APPLICANT: Simmons, Carl R.

; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible

; TITLE OF INVENTION: Proteins and Uses Thereof

; FILE REFERENCE: 35718/242990

; CURRENT APPLICATION NUMBER: US/10/090,035

; CURRENT FILING DATE: 2002-02-28

; PRIOR APPLICATION NUMBER: 60/272,227

; PRIOR FILING DATE: 02/28/2001

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 24

; LENGTH: 92

; TYPE: PRT

; ORGANISM: Trilicium aestivum

; US-10-090-035-24

Query Match 61.5%; Score 32; DB 14; Length 92;

Best Local Similarity 60.0%; Pred. No. 39;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKQAFVFEED 10

Db 33 VKETVQEFED 42

RESULT 8

US-10-099-352-41

; Sequence 41, Application US/10099352

; Publication No. US20030082569A1

; GENERAL INFORMATION:

; APPLICANT: Johnson, Clayton H.

; APPLICANT: York, J. Lyndal

; APPLICANT: McEwen, Joan E.

; TITLE OF INVENTION: Histoplasma Capsulation Catalase Sequences and Their Use in the

; TITLE OF INVENTION: of Histoplasma Capsulation and Histoplasmosis

; FILE REFERENCE: 40715-255988

; CURRENT APPLICATION NUMBER: US/10/099,352

; CURRENT FILING DATE: 2002-03-13

; PRIOR APPLICATION NUMBER: US 60/275,353

; PRIOR FILING DATE: 2001-03-13

; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 41

; LENGTH: 744

; TYPE: PRT

; ORGANISM: Emericella nidulans

; US-10-099-352-41

Query Match 61.5%; Score 32; DB 15; Length 744;

Best Local Similarity 60.0%; Pred. No. 3.2e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKQAFVFEED 10

Db 481 VKKAFSFLD 490

RESULT 9

US-09-864-761-39890

; Sequence 39890, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

; FILE REFERENCE: Aeomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 39890

; LENGTH: 45

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC005940.3

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.8

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.8

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 9.6

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 11

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.2

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 9.5

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 8.5

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.8

; OTHER INFORMATION: EXPRESSED IN SWISSPROT HIT: P51575, EVALUE 8.00e-19

; OTHER INFORMATION: EST_HUMAN HIT: A0118048.1, EVALUE 4.00e-08

US-09-864-761-39890

Query Match 59.6%; Score 31; DB 9; Length 45;

Best Local Similarity 71.4%; Pred. No. 30;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 AFVFEED 10

Db 11 AFLFEYD 17

RESULT 10

US-09-833-082-2

: Sequence 2, Application US/09833082
: Patent No. US20020151480A1
: GENERAL INFORMATION:
: APPLICANT: Chun, Miyoung
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
: TITLE OF INVENTION: CARDIOVASCULAR DISEASE USING 10218
: FILE REFERENCE: MNI-227
: CURRENT APPLICATION NUMBER: US/09/833,082
: CURRENT FILING DATE: 2001-04-10
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 388
: TYPE: PRT
: ORGANISM: Homo Sapiens
US-09-833-082-2

Query Match 59.6%; Score 31; DB 10; Length 388;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 AFVFEFD 10
11:11:11
DB 10 AFLFEYD 16

RESULT 11

US-10-156-761-9506
: Sequence 9506, Application US/10156761
: Publication No. US20030119018A1
: GENERAL INFORMATION:
: APPLICANT: OMURA, SATOSHI
: APPLICANT: IKEDA, HARUO
: APPLICANT: ISHIRAWA, JUN
: APPLICANT: HORIKAWA, HIROSHI
: APPLICANT: SHIBA, TADAYOSHI
: APPLICANT: SAKAKI, YOSHIYUKI
: APPLICANT: HATTORI, MASAHIRA
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-262
: CURRENT APPLICATION NUMBER: US/10/156,761
: CURRENT FILING DATE: 2002-05-29
: PRIOR APPLICATION NUMBER: JP 2001-204089
: PRIOR FILING DATE: 2001-05-30
: PRIOR APPLICATION NUMBER: JP 2001-272697
: PRIOR FILING DATE: 2001-08-02
: NUMBER OF SEQ ID NOS: 15109
: SEQ ID NO 9506
: LENGTH: 431
: TYPE: PRT
: ORGANISM: Streptomyces avermitilis
US-10-156-761-9506

Query Match 59.6%; Score 31; DB 15; Length 431;
Best Local Similarity 70.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKQAFVFEFD 10
111 11 11
DB 43 MKQWRVFEFD 52

RESULT 12

US-09-796-692-995
: Sequence 995, Application US/09796692
: Publication No. US20020198362A1
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Mannion, Paul A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
: TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
: FILE REFERENCE: 2077.001200

: CURRENT APPLICATION NUMBER: US/09/796,692
: CURRENT FILING DATE: 2001-03-01
: PRIOR APPLICATION NUMBER: 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: 60/222,903
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: 60/223,416
: PRIOR FILING DATE: 2000-08-04
: PRIOR APPLICATION NUMBER: 60/223,378
: PRIOR FILING DATE: 2000-08-07
: NUMBER OF SEQ ID NOS: 9597
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 995
: LENGTH: 88
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-796-692-995

Query Match 57.7%; Score 30; DB 10; Length 88;
Best Local Similarity 83.3%; Pred. No. 90;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 FVFEFD 10
1:11111
DB 62 FMFEFD 67

RESULT 13

US-09-796-692-1524
: Sequence 1524, Application US/09796692
: Publication No. US20020198362A1
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Mannion, Paul A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND TE
: TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
: FILE REFERENCE: 2077.001200
: CURRENT APPLICATION NUMBER: US/09/796,692
: CURRENT FILING DATE: 2001-03-01
: PRIOR APPLICATION NUMBER: 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: 60/218,950

; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: 60/222,903
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: 60/223,416
 ; PRIOR FILING DATE: 2000-08-04
 ; PRIOR APPLICATION NUMBER: 60/223,378
 ; PRIOR FILING DATE: 2000-08-07
 ; NUMBER OF SEQ ID NOS: 9597
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1524
 ; LENGTH: 88
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-796-692-1524

Query Match 57.7%; Score 30; DB 10; Length 88;
 Best Local Similarity 83.3%; Pred. No. 90;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 FVFEED 10
 Db 62 FMFEED 67

RESULT 14
 US-10-040-862-995
 ; Sequence 995, Application US/10040862
 ; Publication No. US20030078396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Retter, Marc
 ; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
 ; TITLE OF INVENTION: Hematological Malignancies
 ; FILE REFERENCE: 014058-013520US
 ; CURRENT APPLICATION NUMBER: US/10/040.862
 ; CURRENT FILING DATE: 2001-11-06
 ; PRIOR APPLICATION NUMBER: US 60/186,126
 ; PRIOR FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: US 60/190,479
 ; PRIOR FILING DATE: 2000-03-17
 ; PRIOR APPLICATION NUMBER: US 60/200,545
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: US 60/200,303
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,779
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,999
 ; PRIOR FILING DATE: 2000-05-01
 ; PRIOR APPLICATION NUMBER: US 60/202,084
 ; PRIOR FILING DATE: 2000-05-04
 ; PRIOR APPLICATION NUMBER: US 60/206,201
 ; PRIOR FILING DATE: 2000-05-22
 ; PRIOR APPLICATION NUMBER: US 60/218,950
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 60/222,903
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 60/223,416
 ; PRIOR FILING DATE: 2000-08-04
 ; PRIOR APPLICATION NUMBER: US 60/223,378
 ; PRIOR FILING DATE: 2000-08-07
 ; PRIOR APPLICATION NUMBER: US 09/796,692
 ; NUMBER OF SEQ ID NOS: 10467
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 995
 ; LENGTH: 88
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-040-862-995

Query Match 57.7%; Score 30; DB 15; Length 88;
 Best Local Similarity 83.3%; Pred. No. 90;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 FVFEED 10
 Db 62 FMFEED 67

RESULT 15
 US-10-040-862-1524
 ; Sequence 1524, Application US/10040862
 ; Publication No. US20030078396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Corixa Corporation
 ; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and The
 ; TITLE OF INVENTION: Hematological Malignancies
 ; FILE REFERENCE: 014058-013520US
 ; CURRENT APPLICATION NUMBER: US/10/040.862
 ; CURRENT FILING DATE: 2001-11-06
 ; PRIOR APPLICATION NUMBER: US 60/186,126
 ; PRIOR FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: US 60/190,479
 ; PRIOR FILING DATE: 2000-03-17
 ; PRIOR APPLICATION NUMBER: US 60/200,545
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: US 60/200,303
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,779
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,999
 ; PRIOR FILING DATE: 2000-05-01
 ; PRIOR APPLICATION NUMBER: US 60/202,084
 ; PRIOR FILING DATE: 2000-05-04
 ; PRIOR APPLICATION NUMBER: US 60/206,201
 ; PRIOR FILING DATE: 2000-05-22
 ; PRIOR APPLICATION NUMBER: US 60/218,950
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 60/222,903
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 60/223,416
 ; PRIOR FILING DATE: 2000-08-04
 ; PRIOR APPLICATION NUMBER: US 60/223,378
 ; PRIOR FILING DATE: 2000-08-07
 ; PRIOR APPLICATION NUMBER: US 09/796,692
 ; NUMBER OF SEQ ID NOS: 10467
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1524
 ; LENGTH: 88
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-040-862-1524

Query Match 57.7%; Score 30; DB 15; Length 88;
 Best Local Similarity 83.3%; Pred. No. 90;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 FVFEED 10
 Db 62 FMFEED 67

Search completed: July 24, 2003, 20:21:00
 Job time : 8.28571 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:06:40 ; Search time 0.909091 Seconds
(without alignments)
45.168 Million cell updates/sec

Title: US-09-546-136-1

Perfect score: 52

Sequence: 1 MKQAFVFEFD 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 41799 seqs, 4106219 residues

Total number of hits satisfying chosen parameters: 41799

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	33	63.5	538	5	US-09-820-843B-9
2	31	59.6	345	7	US-60-478-196-3285
3	27	51.9	42	6	US-10-273-573-6148
4	27	51.9	185	1	PCT-US03-20480-34
5	27	51.9	315	6	US-10-275-595A-34
6	27	51.9	406	1	PCT-US02-41612A-448
7	27	51.9	523	6	US-10-451-010-1
8	27	51.9	753	7	US-60-479-073-437
9	27	51.9	854	6	US-10-273-573-6450
10	27	51.9	906	7	US-60-479-073-177
11	27	51.9	906	7	US-60-479-073-178
12	27	51.9	933	6	US-10-372-227-35
13	26	50.0	15	6	US-10-371-525A-137
14	26	50.0	15	6	US-10-371-645-137
15	26	50.0	15	6	US-10-371-260-137
16	26	50.0	15	6	US-10-273-573-6460
17	26	50.0	307	5	US-09-599-760B-2
18	26	50.0	314	6	US-10-331-496A-64
19	26	50.0	332	6	US-10-294-433-660
20	26	50.0	395	6	US-10-323-069A-128
21	26	50.0	501	6	US-10-450-200-3
22	25.5	49.0	343	6	US-10-332-176A-2
23	25.5	49.0	574	6	US-10-332-176A-6
24	25.5	49.0	574	6	US-10-332-176A-8
25	25.5	49.0	599	6	US-10-332-176A-10
26	25	48.1	130	1	PCT-US03-18448-5

27 25 48.1 130 6 US-10-365-112-53
28 25 48.1 179 6 US-10-273-573-6127
29 25 48.1 249 6 US-10-338-785A-3
30 25 48.1 267 6 US-10-273-573-5560
31 25 48.1 362 6 US-10-273-573-6668
32 25 48.1 371 6 US-10-294-433-228
33 25 48.1 436 7 US-60-478-196-3298
34 25 48.1 466 1 PCT-US03-19660-39
35 25 48.1 475 6 US-10-275-296-14
36 25 48.1 482 6 US-10-275-296-13
37 25 48.1 566 6 US-10-273-296-2
38 25 48.1 596 6 US-10-273-573-6495
39 25 48.1 732 6 US-10-273-573-8528
40 25 48.1 900 6 US-10-273-573-8458
41 25 48.1 928 6 US-10-244-081-34
42 25 48.1 1518 1 PCT-US03-06962-40
43 24.5 47.1 380 6 US-10-380-565-28
44 24.5 47.1 520 6 US-10-380-565-35
45 24 46.2 56 6 US-10-273-573-5821

ALIGNMENTS

RESULT 1

US-09-820-843B-9
; Sequence 9, Application US/09820843B
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE
; FILE REFERENCE: PROTEINS USEFUL AS ANTI-INFECTIVES
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843B
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 538
; TYPE: PRT
; ORGANISM: H. Influenzae
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: thiamin ABC transporter, permease protein, putative
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: g11574049
US-09-820-843B-9

Query Match 63.5%; Score 33; DB 5; Length 538;
Best Local Similarity 75.0%; Pred. No. 6.3;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 QAFVFEFD 10

DB 234 QAILPEFD 241

RESULT 2

US-60-478-196-3285
; Sequence 3285, Application US/60478196
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Lemieux, Sebastien
; APPLICANT: Hu, Wenqi
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF ASPERGILLUS FUMIGATUS AI
; FILE REFERENCE: 10182-026-888
; CURRENT APPLICATION NUMBER: US/60/478,196
; CURRENT FILING DATE: 2003-06-13
; NUMBER OF SEQ ID NOS: 4000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3285


```
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
; US-60-478-196-3285

Query Match          59.6%; Score 31; DB 7; Length 345;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 MKQAFVFEED 10
    :| | | :||
Db 196 LKAAFOYQPD 205

RESULT 3
US-10-273-573-6148
; Sequence 6148, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 6148
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-273-573-6148

Query Match          51.9%; Score 27; DB 6; Length 42;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 VFEFD 10
    |||||
Db 5 VFEFD 9

RESULT 4
PCT-US03-20480-34
; Sequence 34, Application PC/TUS0320480
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: ELLIOTT, Vick S.
; APPLICANT: MASON, Patricia M.
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: LEE, Soo Yuen
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: KABLE, Amy E.
; APPLICANT: SHARNAKAR, Anita
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: KHARE, Reena
; APPLICANT: BECHA, Shanya D.
; APPLICANT: MARQUIS, Joseph P.
; APPLICANT: CHIEN, David
; APPLICANT: JIN, Pei
; APPLICANT: CHANG, Hsin-Ru
; APPLICANT: BULLOCH, Sean A.
; APPLICANT: TRAN, Uyen K.
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: ENZYMES
; FILE REFERENCE: PF-1452 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/20480
; CURRENT FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: US 60/393,067
; PRIOR FILING DATE: 2002-06-28

; LENGTH: 345
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
; US-60-478-196-3285

Query Match          59.6%; Score 31; DB 7; Length 345;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 MKQAFVFEED 10
    :| | | :||
Db 196 LKAAFOYQPD 205

PCT-US03-20480-34
; Sequence 34, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 6148
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-273-573-6148

Query Match          51.9%; Score 27; DB 6; Length 42;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 VFEFD 10
    |||||
Db 5 VFEFD 9

PCT-US03-20480-34
; Sequence 34, Application PC/TUS0320480
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: ELLIOTT, Vick S.
; APPLICANT: MASON, Patricia M.
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: LEE, Soo Yuen
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: KABLE, Amy E.
; APPLICANT: SHARNAKAR, Anita
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: KHARE, Reena
; APPLICANT: BECHA, Shanya D.
; APPLICANT: MARQUIS, Joseph P.
; APPLICANT: CHIEN, David
; APPLICANT: JIN, Pei
; APPLICANT: CHANG, Hsin-Ru
; APPLICANT: BULLOCH, Sean A.
; APPLICANT: TRAN, Uyen K.
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: ENZYMES
; FILE REFERENCE: PF-1452 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/20480
; CURRENT FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: US 60/393,067
; PRIOR FILING DATE: 2002-06-28

; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7512863CD1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (185)
; OTHER INFORMATION: unknown or other
PCT-US03-20480-34

Query Match          51.9%; Score 27; DB 1; Length 185;
Best Local Similarity 57.1%; Pred. No. 33;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 4 AFVFEFD 10
    :| | | |
Db 64 SFMFKFD 70

RESULT 5
US-10-275-595A-34
; Sequence 34, Application US/10275595A
; GENERAL INFORMATION:
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LAL, Preeti
; APPLICANT: YAO, Monique G.
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: BATRA, Sajeev
; APPLICANT: KEARNEY, Liam
; APPLICANT: POLICKY, Jennifer L.
; TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0772 USN
; CURRENT APPLICATION NUMBER: US/10/275,595A
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/201,960
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/202,729
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: US 60/209,705
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 60/210,149
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: US 60/213,215
; PRIOR FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PERL Program
; SEQ ID NO 34
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 3035248CD1
US-10-275-595A-34

Query Match          51.9%; Score 27; DB 6; Length 315;
Best Local Similarity 62.5%; Pred. No. 60;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

Qy 1 MKQAFVFE 8
:|||||
Db 193 LKEAFVFE 200

RESULT 6

PCT-US02-41612A-448
: Sequence 448, Application PC/TUS0241612A
: GENERAL INFORMATION:
: APPLICANT: diadexus, Inc.
: APPLICANT: Sun, Yongming
: APPLICANT: Liu, Chenghua
: TITLE OF INVENTION: Compositions and Methods Relating to Endometrial Specific Genes
: FILE REFERENCE: DEX-0378
: CURRENT APPLICATION NUMBER: PCT/US02/41612A
: CURRENT FILING DATE: 2003-06-17
: PRIOR APPLICATION NUMBER: 60/342,756
: PRIOR FILING DATE: 2001-12-21
: NUMBER OF SEQ ID NOS: 600
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 448
: LENGTH: 406
: TYPE: PRT
: ORGANISM: Homo sapien
PCT-US02-41612A-448

Query Match 51.9%; Score 27; DB 1; Length 406;

Best Local Similarity 83.3%; Pred. No. 80;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 AFVFEF 9
:|||||
Db 11 AFAFEF 16

RESULT 7

US-10-451-010-1
: Sequence 1, Application US/10451010
: GENERAL INFORMATION:
: APPLICANT: INCYTE GENOMICS, INC.
: APPLICANT: DUGGAN, Brendan M.
: APPLICANT: XU, Yuming
: APPLICANT: LEE, Ernestine A.
: APPLICANT: LEE, Sally
: APPLICANT: LU, Dyung Aina M.
: APPLICANT: WARREN, Bridget A.
: APPLICANT: YUE, Henry
: APPLICANT: GIETZEN, Kimberly J.
: APPLICANT: HONCHELL, Cynthia D.
: APPLICANT: BURFORD, Neil
: APPLICANT: BAUGHN, Mariah R.
: APPLICANT: TANG, Y. Tom
: APPLICANT: JACKSON, Jennifer L.
: APPLICANT: GANDHI, Aneena R.
: APPLICANT: KALLICK, Deborah A.
: APPLICANT: BANDMAN, Olga
: APPLICANT: GRAUL, Richard C.
: APPLICANT: CHAWLA, Narinder K.
: APPLICANT: LU, Yan
: APPLICANT: RAMKUMAR, Jayalaxmi
: APPLICANT: YAO, Monique G.
: APPLICANT: LAL, Preeti G.
: TITLE OF INVENTION: CELL ADHESION PROTEINS
: FILE REFERENCE: PF-0867 USN
: CURRENT APPLICATION NUMBER: US/10/451,010
: CURRENT FILING DATE: 2003-06-17
: PRIOR APPLICATION NUMBER: PCT/US01/49206
: PRIOR FILING DATE: 2001-12-18
: PRIOR APPLICATION NUMBER: US 60/256,542
: PRIOR FILING DATE: 2000-12-18
: PRIOR APPLICATION NUMBER: US 60/259,604
: PRIOR FILING DATE: 2000-12-22

: PRIOR APPLICATION NUMBER: US 60/260,101
: PRIOR FILING DATE: 2001-01-05
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: PERL Program
: SEQ ID NO 1
: LENGTH: 523
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 4350981CD1
US-10-451-010-1

Query Match 51.9%; Score 27; DB 6; Length 523;

Best Local Similarity 62.5%; Pred. No. 1.1e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KQAFVFEF 9
:|||||
Db 482 EQAFLVEF 489

RESULT 8

US-60-479-073-437
: Sequence 437, Application US/60479073
: GENERAL INFORMATION:
: APPLICANT: De Wilde, Gert Jules Hector
: APPLICANT: Saunders, Michael John Scott
: APPLICANT: Logghe, Marc Georges
: TITLE OF INVENTION: Amino acid sequences useful for developing compounds for the
: TITLE OF INVENTION: prevention and/or treatment of metabolic diseases and nucleic
: TITLE OF INVENTION: sequences encoding such amino acid sequences.
: FILE REFERENCE: D00590.70042.US
: CURRENT APPLICATION NUMBER: US/60/479,073
: CURRENT FILING DATE: 2003-06-17
: NUMBER OF SEQ ID NOS: 526
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 437
: LENGTH: 753
: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-479-073-437

Query Match 51.9%; Score 27; DB 7; Length 753;

Best Local Similarity 50.0%; Pred. No. 1.6e+02;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KQAFVFEF 9
:|||||
Db 737 KEGTFEY 744

RESULT 9

US-10-273-573-6450
: Sequence 6450, Application US/10273573
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc
: TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
: FILE REFERENCE: 21272-066
: CURRENT APPLICATION NUMBER: US/10/273,573
: CURRENT FILING DATE: 2002-10-18
: PRIOR APPLICATION NUMBER: 09/522,929
: PRIOR FILING DATE: 2000-04-18
: PRIOR APPLICATION NUMBER: 09/770,160
: PRIOR FILING DATE: 2001-01-26
: NUMBER OF SEQ ID NOS: 10994
: SOFTWARE: Custom
: SEQ ID NO 6450
: LENGTH: 854
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: DOMAIN

```

: LOCATION: (431)...(481)
: OTHER INFORMATION: Eukaryotic DNA topoisomerase I proteins domain identified by
: OTHER INFORMATION: EMATRIX, accession number BL00176A, p-value=1.000e-40, raw score
: OTHER INFORMATION: 28.37
: FEATURE:
: NAME/KEY: DOMAIN
: LOCATION: (204)...(854)
: OTHER INFORMATION: Eukaryotic DNA topoisomerase I domain identified by PFam,
: OTHER INFORMATION: accession name Topoisomerase_I, E-value=0, PFam score of 1431.8
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (1)...(854)
: OTHER INFORMATION: Xaa - X or * as defined in Table 2
US-10-273-573-6450

```

```

Query Match      51.9%  Score 27; DB 6; Length 854;
Best Local Similarity 62.5%  Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      3 QAFVFEPD 10
Db      609 QEVVVEPD 616

```

RESULT 10

```

US-60-479-073-177
: Sequence 177, Application US/60479073.
: GENERAL INFORMATION:
: APPLICANT: De Wilde, Gert Jules Hector
: APPLICANT: Saunders, Michael John Scott
: APPLICANT: Logghe, Marc Georges
: TITLE OF INVENTION: Amino acid sequences useful for developing compounds for the
: TITLE OF INVENTION: prevention and/or treatment of metabolic diseases and nucleotide
: FILE REFERENCE: D00590.70042.US
: CURRENT APPLICATION NUMBER: US/60/479,073
: CURRENT FILING DATE: 2003-06-17
: NUMBER OF SEQ ID NOS: 526
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 177
: LENGTH: 906
: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-479-073-177

```

```

Query Match      51.9%  Score 27; DB 7; Length 906;
Best Local Similarity 62.5%  Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 MKQAFVFE 8
Db      801 LKEAFVVE 808

```

RESULT 11

```

US-60-479-073-178
: Sequence 178, Application US/60479073
: GENERAL INFORMATION:
: APPLICANT: De Wilde, Gert Jules Hector
: APPLICANT: Saunders, Michael John Scott
: APPLICANT: Logghe, Marc Georges
: TITLE OF INVENTION: Amino acid sequences useful for developing compounds for the
: TITLE OF INVENTION: prevention and/or treatment of metabolic diseases and nucleotide
: FILE REFERENCE: D00590.70042.US
: CURRENT APPLICATION NUMBER: US/60/479,073
: CURRENT FILING DATE: 2003-06-17
: NUMBER OF SEQ ID NOS: 526
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 178
: LENGTH: 906
: TYPE: PRT
: ORGANISM: Homo sapiens

```

US-60-479-073-178

```

Query Match      51.9%  Score 27; DB 7; Length 906;
Best Local Similarity 62.5%  Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 MKQAFVFE 8
Db      801 LKEAFVVE 808

```

RESULT 12

```

US-10-372-227-35
: Sequence 35, Application US/10372227
: GENERAL INFORMATION:
: APPLICANT: Hitachi, Ltd.
: TITLE OF INVENTION: OPAL-INTERACTING PROTEINS AND USE THEREOF
: FILE REFERENCE: HITA.0168
: CURRENT APPLICATION NUMBER: US/10/372,227
: CURRENT FILING DATE: 2003-02-25
: PRIOR APPLICATION NUMBER: US 60/359,374
: PRIOR FILING DATE: 2002-02-26
: PRIOR APPLICATION NUMBER: US 60/407,247
: PRIOR FILING DATE: 2002-09-03
: NUMBER OF SEQ ID NOS: 45
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 35
: LENGTH: 933
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-372-227-35

```

```

Query Match      51.9%  Score 27; DB 6; Length 933;
Best Local Similarity 83.3%  Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      2 KQAFVF 7
Db      824 ROAFVF 829

```

RESULT 13

```

US-10-371-525A-137
: Sequence 137, Application US/10371525A
: GENERAL INFORMATION:
: APPLICANT: EPIMUNE Inc.
: APPLICANT: Fikes, John D.
: APPLICANT: Hermanson, Gary G.
: APPLICANT: Sette, Alessandro
: APPLICANT: Ishioka, Glenn Y.
: APPLICANT: Livingston, Brian
: APPLICANT: Chesnut, Robert W.
: APPLICANT: Epimmune Inc.
: TITLE OF INVENTION: Expression Vectors for Stimulating an
: TITLE OF INVENTION: Immune Response and Methods of Using the Same
: FILE REFERENCE: 39963-20022.12
: CURRENT APPLICATION NUMBER: US/10/371,525A
: CURRENT FILING DATE: 2003-05-16
: PRIOR APPLICATION NUMBER: US 09/078,904
: PRIOR FILING DATE: 1998-05-13
: PRIOR APPLICATION NUMBER: US 60/085,751
: PRIOR FILING DATE: 1998-05-15
: NUMBER OF SEQ ID NOS: 463
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 137
: LENGTH: 15
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: HBV POL 661 (peptide 1298-06)
US-10-371-525A-137

```

```

Query Match      50.0%  Score 26; DB 6; Length 15;

```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HBV POL 661 (peptide 1298.06)
US-10-371-260-137

Query Match      50.0%; Score 26; DB 6; Length 15;
Best Local Similarity 83.3%; Pred. No. 3.3;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 KOAEVF 7
      11111
Db      1 KOAEVF 6

Search completed: July 24, 2003, 20:19:44
Job time : 1.90909 secs

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HBV POL 661 (peptide 1298.06)
US-10-371-645-137

Sequence 137, Application US/10371645
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.11
; CURRENT APPLICATION NUMBER: US/10/371,645
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 137
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HBV POL 661 (peptide 1298.06)
US-10-371-645-137

Query Match      50.0%; Score 26; DB 6; Length 15;
Best Local Similarity 83.3%; Pred. No. 3.3;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 KOAEVF 7
      11111
Db      1 KOAEVF 6

Sequence 137, Application US/10371260
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.13
; CURRENT APPLICATION NUMBER: US/10/371,260
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 137
; LENGTH: 15
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:04:00 ; Search time 3.24675 Seconds
(without alignments)
296.200 Million cell updates/sec

Title: US-09-546-136-1

Perfect score: 52

Sequence: 1 MKQAFVFEFD 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	36	69.2	271	JQ2285	nodulin-26 - soybe
2	36	69.2	271	JQ2286	nodulin-26 - soybe
3	36	69.2	271	S01444	nodulin-26 precurs
4	34	65.4	285	B90445	regucalcin homolog
5	34	65.4	339	T08677	hypothetical prote
6	33	63.5	431	RGECEP	sensor kinase (EC
7	33	63.5	431	B44753	hypothetical prote
8	33	63.5	431	B90689	hypothetical prote
9	33	63.5	431	F85535	hypothetical prote
10	33	63.5	538	D64164	hypothetical prote
11	33	63.5	555	F84648	hypothetical prote
12	32	61.5	45	PC4177	phosphoglycolate p
13	32	61.5	84	A27144	larval serum prote
14	32	61.5	100	B27144	larval serum prote
15	32	61.5	141	B69385	hypothetical prote
16	32	61.5	167	T37486	probable signal pe
17	32	61.5	188	E71495	probable peptidogl
18	32	61.5	202	G81653	peptidoglycan asso
19	32	61.5	371	T49786	hypothetical prote
20	32	61.5	395	F64351	hypothetical prote
21	32	61.5	408	C69619	ABC transporter (m
22	32	61.5	552	H86390	TIK7.21 protein -
23	32	61.5	744	S68115	catalase (EC 1.11.
24	32	61.5	804	YFBSB	phenylalanine-trna
25	32	61.5	842	S49124	carbon-monoxide de
26	31	59.6	124	AC0803	probable bacteriop
27	31	59.6	230	F71122	hypothetical prote
28	31	59.6	283	S64347	hypothetical prote
29	31	59.6	329	D96834	hypothetical prote

30	31	59.6	368	2	T27432	hypothetical prote
31	31	59.6	385	2	T23065	hypothetical prote
32	31	59.6	396	2	E70233	hypothetical prote
33	31	59.6	399	2	S71927	ATP-gated ion chan
34	31	59.6	439	2	T45798	hypothetical prote
35	31	59.6	445	2	T32366	hypothetical prote
36	31	59.6	449	2	C95332	probable copper-co
37	31	59.6	449	2	A98239	hypothetical prote
38	31	59.6	449	2	AC3047	multicopper oxidas
39	31	59.6	456	1	C64772	probable transport
40	31	59.6	456	2	A90689	probable transport
41	31	59.6	456	2	E85539	probable transport
42	31	59.6	459	2	T10307	hypothetical prote
43	31	59.6	471	1	S62590	peptidyl-prolyl ci
44	31	59.6	569	2	T38627	hypothetical dna b
45	31	59.6	587	2	T02829	conserved hypothet

ALIGNMENTS

RESULT 1
JQ2285
nodulin-26 - soybean
C:Species: Glycine max (soybean)
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Mar-1999
C:Accession: JQ2285
R:Miao, G.H.; Verma, D.P.S.
Plant Cell 5, 781-794, 1993
A:Title: Soybean nodulin-26 gene encoding a channel protein is expressed only in the
A:Reference number: JQ2285; MUID:93372569; PMID:7689881
A:Accession: JQ2285
A:Molecule type: DNA
A:Residues: 1-271 <MIA>
C:Comment: The protein is a major peribacteroid membrane protein in soybean root nod
C:Genetics:
A:Introns: 39/3; 114/3; 180/3; 201/2
C:Superfamily: nodulin-26
C:Keywords: transmembrane protein
F:40-60/Domain: transmembrane #status predicted <TM1>
F:67-88/Domain: transmembrane #status predicted <TM2>
F:115-133/Domain: transmembrane #status predicted <TM3>
F:156-175/Domain: transmembrane #status predicted <TM4>
F:183-203/Domain: transmembrane #status predicted <TM5>
F:227-244/Domain: transmembrane #status predicted <TM6>

Query Match 69.2%; Score 36; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QAFVFEF 9
IIIIIII
Db 155 QAFVFEF 161

RESULT 2
JQ2286
nodulin-26 - soybean
N:Alternate names: channel protein
C:Species: Glycine max (soybean)
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Mar-1999
C:Accession: JQ2286
R:Miao, G.H.; Verma, D.P.S.
Plant Cell 5, 781-794, 1993
A:Title: Soybean nodulin-26 gene encoding a channel protein is expressed only in the
A:Reference number: JQ2285; MUID:93372569; PMID:7689881
A:Accession: JQ2286
A:Molecule type: mRNA
A:Residues: 1-271 <MIA>
A:Note: the authors translated the codon GTT for residue 211 as Ala
C:Comment: This protein is a major peribacteroid membrane protein in soybean root nc
C:Superfamily: nodulin-26

Query Match 69.2%; Score 36; DB 2; Length 271;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QAFVFEF 9
 |||||

DB 155 QAFVFEF 161

RESULT 3

S01444
 nodulin-26 precursor - soybean
 C:Species: Glycine max (soybean)
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 17-Nov-2000
 C:Accession: S01444; A26445
 R:Sandal, N.N.; Marcker, K.A.
 Nucleic Acids Res. 16, 9347, 1988
 A:Title: Soybean nodulin 26 is homologous to the major intrinsic protein of the bovine lens
 A:Reference number: S01444; MUID:89016643; PMID:3174457
 A:Accession: S01444
 A:Molecule type: mRNA
 A:Residues: 1-271 <SAN>
 A:Cross-references: EMBL:X12659
 R:Fortin, M.G.; Morrison, N.A.; Verma, D.P.S.
 Nucleic Acids Res. 15, 813-824, 1987
 A:Title: Nodulin-26, a peribacteroid membrane nodulin is expressed independently of the
 A:Reference number: A26445; MUID:87146395; PMID:3822816
 A:Accession: A26445
 A:Molecule type: mRNA
 A:Residues: 67-271 <FOR>
 A:Experimental source: seed, cv. Prize
 A:Superfamily: nodulin-26
 C:Keywords: phosphoprotein; transmembrane protein

Query Match 69.2%; Score 36; DB 2; Length 271;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QAFVFEF 9
 |||||

DB 155 QAFVFEF 161

RESULT 4

B90445
 regucalcin homolog [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
 C:Accession: B90445
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: B90445
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-285 <KUR>
 A:Cross-references: GB:AE006641; MUID:q13816023; PIDN:AAK42817.1; GSPDB:GNO0155
 C:Genetics:
 A:Gene: SSO2705
 C:Superfamily: senescence marker protein-30

Query Match 65.4%; Score 34; DB 2; Length 285;
 Best Local Similarity 50.0%; Pred. No. 26;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKQAFVFEF 10
 ::|||

DB 165 VRKVFVFEF 174

RESULT 5

T08677
 hypothetical protein DKF2p564G2222.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 23-Jul-1999
 C:Accession: T08677
 R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, May 1999
 A:Reference number: Z16469
 A:Accession: T08677
 A:Molecule type: mRNA
 A:Residues: 1-339 <WAM>
 A:Cross-references: EMBL:AL049945
 A:Experimental source: fetal brain; clone DKF2p564G2222
 C:Genetics:
 A:Note: DKF2p564G2222.1

Query Match 65.4%; Score 34; DB 2; Length 339;
 Best Local Similarity 60.0%; Pred. No. 31;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKQAFVFEF 10
 ::|||

DB 280 LKVLFLFEF 289

RESULT 6

RGECFR
 sensor kinase (EC 2.7.3.-) phoR - Escherichia coli (strain K-12)
 N:Alternate names: phosphate regulon sensor protein phoR
 C:Species: Escherichia coli
 C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 01-Mar-2002
 C:Accession: A25557; H64768; S11888
 R:Makino, K.; Shinagawa, H.; Amemura, M.; Nakata, A.
 J. Mol. Biol. 192, 549-556, 1986
 A:Title: Nucleotide sequence of the phoR gene, a regulatory gene for the phosphate
 A:Reference number: A25557; MUID:87169739; PMID:3550103
 A:Accession: A25557
 A:Molecule type: DNA
 A:Residues: 1-431 <MAK>
 A:Cross-references: GB:X04704; MUID:q42393; PIDN:CAA28409.1; PID:q581188
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.
 .A.; Rose, D.J.; Mau, B.; Shaoh, Y.
 Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: H64768
 A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-431 <BLAT>

A:Cross-references: GB:AE000146; GB:U00096; MUID:g1786596; PIDN:AACT3503.1; PID:g17E

A:Experimental source: strain K-12, substrain MG1655

R:Yamada, M.; Makino, K.; Shinagawa, H.; Nakata, A.

Mol. Gen. Genet. 220, 366-372, 1990

A:Title: Regulation of the phosphate regulon of Escherichia coli: properties of phc

A:Reference number: S11888; MUID:90251245; PMID:2187152

A:Accession: S11888

A:Molecule type: DNA

A:Residues: 1-13 <YAM>

C:Genetics:

A:Gene: phoR

A:Map position: 9 min

A:Start codon: GTG

C:Function:

A:Description: transcription regulation; involved in transcription activation of th

A:Note: phosphorylated phoR protein phosphorylates phoB protein; phosphorylated phc

C:Superfamily: phosphate regulon regulatory protein; sensor histidine kinase homolo

C:Keywords: autophosphorylation; membrane protein; phosphate transport; phosphohist

F:10-26/Domain: transmembrane #status predicted <TM1>

F:28-44/Domain: transmembrane #status predicted <TM2>

F:160-422/Domain: sensor histidine kinase homology <SHK>

F:213/Binding site: phosphate (His) (covalent) #status predicted

F:213/Binding site: phosphate (His) (covalent) (by autophosphorylation) #status predicted

Query Match 63.5%; Score 33; DB 1; Length 431;
 Best Local Similarity 66.7%; Pred. No. 64;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KOAFVFEED 10
 II I I I I
 DB 298 KOTTFEED 306

RESULT 7
 B44753
 sensor kinase (EC 2.7.3.-) phor - Shigella dysenteriae
 C:Species: Shigella dysenteriae
 C:Date: 03-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 04-Feb-2000
 C:Accession: B44753
 R:Lee, T.Y.; Makino, K.; Shinagawa, H.; Anemura, M.; Nakata, A.
 J. Bacteriol. 171, 6593-6599, 1989
 A:Title: Phosphate regulon in members of the family Enterobacteriaceae: comparison of the
 A:Reference number: A44753; MUID:90078103; PMID:2556368
 A:Accession: B44753
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-431 <LEE>
 A:Cross-references: EMBL:M31793; NID:gl252772; PIDN:AAA26536.1; PID:g294893
 A:Note: the authors translated the codon GCG for residue 239 as Ala
 C:Superfamily: phosphate regulon regulatory protein; sensor histidine kinase homology
 C:Keywords: autophosphorylation; phosphohistidine; phosphoprotein; phosphotransferase
 F:186-422/domain: sensor histidine kinase homology <SHK>
 F:213/Binding site: phosphate (His) (covalent) (by autophosphorylation) #status predicted

Query Match 63.5%; Score 33; DB 2; Length 431;
 Best Local Similarity 66.7%; Pred. No. 64;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KOAFVFEED 10
 II I I I I
 DB 298 KOTTFEED 306

RESULT 8
 B90685
 hypothetical protein ECs0450 [imported] - Escherichia coli (strain O157:H7, substrain R1
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: B90685
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
 A:Reference number: A99629; MUID:21156731; PMID:11258796
 A:Accession: B90685
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-431 <HAY>
 A:Cross-references: PIDN:BA33873.1; PID:gl3359907; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain R1MD 050952
 C:Genetics:
 A:Gene: ECs0450
 C:Superfamily: phosphate regulon regulatory protein; sensor histidine kinase homology

Query Match 63.5%; Score 33; DB 2; Length 431;
 Best Local Similarity 66.7%; Pred. No. 64;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KOAFVFEED 10
 II I I I I
 DB 298 KOTTFEED 306

RESULT 9
 F85535
 sensor kinase (EC 2.7.3.-) phor - Shigella dysenteriae
 C:Species: Shigella dysenteriae
 C:Date: 03-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 04-Feb-2000
 C:Accession: B44753
 R:Lee, T.Y.; Makino, K.; Shinagawa, H.; Anemura, M.; Nakata, A.
 J. Bacteriol. 171, 6593-6599, 1989
 A:Title: Phosphate regulon in members of the family Enterobacteriaceae: comparison of the
 A:Reference number: A44753; MUID:90078103; PMID:2556368
 A:Accession: B44753
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-431 <LEE>
 A:Cross-references: EMBL:M31793; NID:gl252772; PIDN:AAA26536.1; PID:g294893
 A:Note: the authors translated the codon GCG for residue 239 as Ala
 C:Superfamily: phosphate regulon regulatory protein; sensor histidine kinase homology
 C:Keywords: autophosphorylation; phosphohistidine; phosphoprotein; phosphotransferase
 F:186-422/domain: sensor histidine kinase homology <SHK>
 F:213/Binding site: phosphate (His) (covalent) (by autophosphorylation) #status predicted

hypothetical protein phor [imported] - Escherichia coli (strain O157:H7, substrain E
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: F85535
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Ma
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apod
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: F85535
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-431 <STO>
 A:Cross-references: GB:AE005174; NID:gl2513246; PIDN:AAG54746.1; GSPDB:GN00145; UWGF
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: phor
 C:Superfamily: phosphate regulon regulatory protein; sensor histidine kinase homology

Query Match 63.5%; Score 33; DB 2; Length 431;
 Best Local Similarity 66.7%; Pred. No. 64;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KOAFVFEED 10
 II I I I I
 DB 298 KOTTFEED 306

RESULT 10
 D64164
 hypothetical protein H1020 - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: D64164
 R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavag
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidma
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.N
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vent
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: D64164
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-538 <TIGR>
 A:Cross-references: GB:U32782; GB:I42023; NID:gl574041; PIDN:ACC22679.1; PID:gl57404
 A:Note: best homolog was a hypothetical protein from Escherichia coli
 C:Superfamily: sfub protein

Query Match 63.5%; Score 33; DB 1; Length 538;
 Best Local Similarity 75.0%; Pred. No. 80;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 QAFVFEED 10
 II I I I I
 DB 234 QAILFEED 241

RESULT 11
 F84648
 hypothetical protein At2g25460 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: F84648
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallor
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Vent
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: F84648
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-555 <STO>

A:Cross-references: GB:AE002093; NID:q4432857; PIDN:AD20705.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g25460

A:Map position: 2

Query Match 61.5%; Score 33; DB 2; Length 555;
Best Local Similarity 55.6%; Pred. No. 82;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 KQAFVFED 10

|||:: ||

Db 476 KQAYILKFD 484

RESULT 12

PC4177

Phosphoglycolate phosphatase (EC 3.1.3.18) - *Synechococcus* sp. (fragment)

C:Species: *Synechococcus* sp.

C:Date: 15-Aug-1996 #sequence_revision 18-Oct-1996 #text_change 07-May-1999

C:Accession: PC4177

R:Liu, Y.; Tsinoiremas, N.F.

Gene 172, 105-109, 1996

A:Title: An unusual gene arrangement for the putative chromosome replication origin and

A:Reference number: PC4177; MUID:96257262; PMID:8654968

A:Contents: PCC7942

A:Accession: PC4177

A:Molecule type: DNA

A:Residues: 1-45 <LI0>

A:Cross-references: GB:U33322

C:Genetics:

A:Gene: cbb2p

A:Start codon: TTG

C:Keywords: phosphoric monoester hydrolase

Query Match 61.5%; Score 32; DB 2; Length 45;
Best Local Similarity 62.5%; Pred. No. 10;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QAFVFED 10

||:: |||

Db 2 QAIIFD 9

RESULT 13

A27144

larval serum protein 1 alpha chain precursor - fruit fly (*Drosophila melanogaster*) (frag

C:Species: *Drosophila melanogaster*

C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 21-Jun-2002

C:Accession: A27144

R:Delaney, S.J.; Smith, D.F.; McClelland, A.; Sunkel, C.; Glover, D.M.

J. Mol. Biol. 189, 1-11, 1986

A:Title: Sequence conservation around the 5' ends of the larval serum protein 1 genes of

A:Reference number: A92926; MUID:87060914; PMID:3097321

A:Accession: A27144

A:Molecule type: DNA

A:Residues: 1-84

C:Genetics:

A:Gene: FlyBase:lspl-alpha

A:Cross-references: FlyBase:FBgn0002562

C:Superfamily: arylphorin

Query Match 61.5%; Score 32; DB 2; Length 84;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKQAFVFE 8

|||:: ||

Db 34 MKQKFLFE 41

RESULT 14

B27144

larval serum protein 1 beta chain precursor - fruit fly (*Drosophila melanogaster*)

C:Species: *Drosophila melanogaster*

C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 21-Jun-2002

C:Accession: B27144

R:Delaney, S.J.; Smith, D.F.; McClelland, A.; Sunkel, C.; Glover, D.M.

J. Mol. Biol. 189, 1-11, 1986

A:Title: Sequence conservation around the 5' ends of the larval serum protein 1 ge

A:Reference number: A92926; MUID:87060914; PMID:3097321

A:Accession: B27144

A:Molecule type: DNA

A:Residues: 1-100

A:Cross-references: GB:X03873; NID:98193; PIDN:CAA27507.1; PID:98194

C:Genetics:

A:Gene: FlyBase:lspl-beta

A:Cross-references: FlyBase:FBgn0002563

C:Superfamily: arylphorin

Query Match 61.5%; Score 32; DB 2; Length 100;
Best Local Similarity 75.0%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKQAFVFE 8

|||:: ||

Db 34 MKQKFLFE 41

RESULT 15

H69385

hypothetical protein AF1089 - *Archaeoglobus fulgidus*

C:Species: *Archaeoglobus fulgidus*

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000

C:Accession: H69385

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; L

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness,

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Syke

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing ar

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: H69385

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-141 <KLE>

A:Cross-references: GB:AE001028; GB:AE000782; NID:92689351; PIDN:AAB90162.1; PID:92

C:Superfamily: *Archaeoglobus fulgidus* conserved hypothetical protein AF1711

Query Match 61.5%; Score 32; DB 1; Length 141;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 FVFEFD 10

||:: |||

Db 14 FVFEFD 19

Search completed: July 24, 2003, 20:10:22

Job time : 6.24675 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 19:47:29 ; Search time 1.75325 seconds
(without alignments)
268.226 Million cell updates/sec

Title: US-09-546-136-1
Perfect score: 52
Sequence: 1 MKQAFVFEED 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	69.2	271	1 NO26_SOYBN	P08995 glycine max
2	36	69.2	749	1 CATA_ASPEU	P78574 aspergillus
3	34	65.4	1332	1 IKAP_HUMAN	O95163 homo sapien
4	33	63.5	431	1 PHOR_ECOLI	P08400 escherichia
5	33	63.5	431	1 PHOR_SHIDY	P45609 shigella dy
6	33	63.5	538	1 THIP_HAEN	P44985 haemophilus
7	32	61.5	212	1 GPH_SYN7	Q55039 synechococc
8	32	61.5	395	1 Y414_METJA	Q57857 methanococc
9	32	61.5	408	1 ECSB_BACSU	P55340 bacillus su
10	32	61.5	744	1 CATA_EMENI	P53305 emericeella
11	32	61.5	789	1 LPIA_DROME	P11995 drosophila
12	32	61.5	789	1 LPIB_DROME	P11996 drosophila
13	32	61.5	804	1 SYFB_BACSU	P17922 bacillus su
14	32	61.5	1297	1 PUR4_PASMU	Q961w4 pasteurella
15	31	59.6	283	1 YG21_YEAST	P53234 saccharomyc
16	31	59.6	368	1 THIL_SULTO	O96ve5 sulfolobus
17	31	59.6	388	1 P2X4_HUMAN	Q99571 homo sapien
18	31	59.6	399	1 P2X1_HUMAN	P51575 homo sapien
19	31	59.6	454	1 YAJR_ECOLI	P77726 escherichia
20	31	59.6	459	1 O30_NPVOX	O10293 orgyia pseu
21	31	59.6	471	1 YAL5_SCHPO	Q09928 schizosacch
22	31	59.6	548	1 SVE_THEVO	O979q0 thermoplasm
23	31	59.6	2561	1 PPS1_BACSU	P39845 bacillus su
24	30	57.7	127	1 CBIW_BACME	O87688 bacillus me
25	30	57.7	220	1 YHCW_BACSU	P54607 bacillus su
26	30	57.7	230	1 CRP_RAT	P48199 rattus norv
27	30	57.7	241	1 RDKA_HAEN	O86224 haemophilus
28	30	57.7	260	1 HA2Q_HUMAN	P20036 homo sapien
29	30	57.7	396	1 YC44_CYACA	O19913 cyanidium c
30	30	57.7	408	1 MTVL_VIBS3	Q03055 vibrio sp.
31	30	57.7	419	1 SECY_PAVLU	P28540 pavlova lut
32	30	57.7	426	1 RMS5_NEUCR	P23351 neurospora
33	30	57.7	452	1 PLM1_PLAFA	P39898 plasmodium

34 30 57.7 574 1 CDAS_THEET P29964 thermocanaer
35 30 57.7 808 1 SYFB_BACHD O94896 bacillus ha
36 30 57.7 829 1 TOP1_XENLA P41512 xenopus lae
37 30 57.7 843 1 DPOL_HPBVT Q05486 hepatitis b
38 30 57.7 845 1 SCPI_MESAU Q60563 mesocricetu
39 30 57.7 851 1 CE05_MOUSE Q82h3 mus musculu
40 30 57.7 997 1 SCPI_RAT Q03410 rattus norv
41 30 57.7 1107 1 YLK2_CAEEL P41950 caenorhabdi
42 30 57.7 1286 1 YKV5_YEAST P28273 saccharomyc
43 30 57.7 1435 1 DFO3_MYCPU P47729 mycoplasma
44 30 57.7 1451 1 AT7B_RAT Q64535 rattus norv
45 30 57.7 1462 1 AT7B_MOUSE Q64446 mus musculu

ALIGNMENTS

RESULT 1
NO26_SOYBN STANDARD; PRT; 271 AA.
AC P08995;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Nodulin-26 (N-26).
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Evans;
RX MEDLINE=89018643; PubMed=3174457;
RA Sandall N.N., Marcker K.A.;
RT "Soybean nodulin 26 is homologous to the major intrinsic protein of
the bovine lens fiber membrane.";
RL Nucleic Acids Res. 16:9347-9347(1988).
RN [2]
RP SEQUENCE OF 67-271 FROM N.A.
RC STRAIN=cv. Prize;
RX MEDLINE=87146395; PubMed=3822816;
RA Fortin M.G., Morrison N.A., Verma D.P.S.;
RT "Nodulin-26, a peribacteroid membrane nodulin is expressed
independently of the development of the peribacteroid compartment.";
RL Nucleic Acids Res. 15:813-824(1987).
RN [3]
RP REVISIONS TO 184 AND 257.
RC STRAIN=cv. Prize;
RA Miao G.H.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [4]
RP PHOSPHORYLATION BY CDPK.
RX MEDLINE=93003032; PubMed=1390682;
RA Waever C.D., Roberts D.M.;
RT "Determination of the site of phosphorylation of nodulin 26 by the
calcium-dependent protein kinase from soybean nodules.";
RL Biochemistry 31:8954-8959(1992).
CC -!- FUNCTION: NOT KNOWN, MAY FUNCTION IN TRANSPORTING SMALL
MOLECULES ACROSS THE PERIBACTEROID MEMBRANES (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PERIBACTEROID
MEMBRANE.
CC -!- INDUCTION: DURING NODULATION IN LEGUME ROOTS AFTER RHIZOBIUM
INFECTION.
CC -!- SIMILARITY: BELONGS TO THE MIP/AQUAPORIN FAMILY (TC 1.A.8).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).
or send an email to license@isb-sib.ch.

```

CC EMBL: X04782; CAA28471.1; -
CC EMBL: X12659; CAA31186.1; -
CC HSP: P11244; IFX8.
CC InterPro: IPR000425; MIP_family.
CC Pfam: PF00230; MIP; 1.
CC PRINTS: PR00783; MINTRINSICP.
CC PRODOM: PD000295; MIP_family; 1.
CC TIGRFAMS: TIGR00861; MIP; 1.
CC PROSITE: PS00221; MIP; 1.
CC Modulation; Nitrogen fixation; Transmembrane; Phosphorylation.
KW Nodulation; Nitrogen fixation; Transmembrane; Phosphorylation.
FT TRANSMEM 154 173 POTENTIAL.
FT MOD_RES 262 262 PHOSPHORYLATION (BY CDPK).
SQ SEQUENCE 271 AA; 28935 MW; EA323421D39042B4 CRC64;

Query Match 69.28; Score 36; DB 1; Length 271;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QAFVFEF 9
Db 155 QAFVFEF 161
IIIIIII

RESULT 2
CATA_AS PFU STANDARD; PRT; 749 AA.
AC P78574;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Catalase A (EC 1.11.1.6).
GN CATA.
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
ON NCBI_TaxID=5085;
[1]
RP SEQUENCE FROM N.A.
RA Wysocki D.R., Diamond R.D., Robbins P.W.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Occurs in almost all aerobically respiring organisms and
CC serves to protect cells from the toxic effects of hydrogen
CC peroxide.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP.
CC -!- SIMILARITY: Belongs to the catalase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U87630; AB47761.1; -
CC HSP: P21179; ICF9.
CC InterPro: IPR002226; Catalase.
CC Pfam: PF00199; catalase; 1.
CC PRINTS: PR00067; CATALASE.
CC PRODOM: PD000510; Catalase; 1.
CC PROSITE: PS00437; CATALASE_1; 1.
CC PROSITE: PS00438; CATALASE_2; 1.
CC Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide.
KW ACT_SITE 93 93 BY SIMILARITY.
FT ACT_SITE 166 166 BY SIMILARITY.
FT METAL 380 380 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 749 AA; 84550 MW; 298854CFB4C826E0 CRC64;

Query Match 69.28; Score 36; DB 1; Length 749;
Best Local Similarity 70.0%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 MKQAFVFEFD 10
Db 487 MKKAFSPELD 496
IIIIIII

RESULT 3
IKAP_HUMAN STANDARD; PRT; 1332 AA.
ID IKAP_HUMAN
AC O95163; Q9H327; Q9UG87;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE IkappaB kinase complex-associated protein (IKK complex-associated
DE protein) (p150).
DE protein) (p150).
GN IKKAP OR IKAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craydata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP TISSUE=Cervical carcinoma;
RX MEDLINE=98421679; Pubmed=9751059;
RA Cohen L., Henzel W.J., Baeuerle P.A.;
RT "IKAP is a scaffold protein of the IkappaB kinase complex.";
RL Nature 395:292-296(1998).
[2]
RN SEQUENCE FROM N.A., AND VARIANT FD PRO-696.
RX MEDLINE=21090531; Pubmed=11179008;
RA Staugenhaupt S.A., Blumenfeld A., Gill S.P., Leyne M., Mull J.,
RA Cuaungco M.P., Liebert C.B., Chadwick B.P., Idelson M., Reznik L.,
RA Robbins C.M., Makalowska I., Brownstein M.J., Krappmann D.,
RA Scheideit C., Maayan C., Axelrod F.B., Gusella J.F.;
RT "Tissue-specific expression of a splicing mutation in the IKKAP gene
RT causes familial dysautonomia.";
RL Am. J. Hum. Genet. 68:598-605(2001).
[3]
RN SEQUENCE OF 961-1332 FROM N.A.
RP TISSUE=Brain;
RA Wambutt R., Heubner D., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
[4]
RN VARIANT FD PRO-696, AND EFFECT ON PHOSPHORYLATION.
RX MEDLINE=21090544; Pubmed=11179021;
RA Anderson S.L., Coli R., Daly I.W., Kichula E.A., Rork M.J.,
RA Volpi S.A., Ekstein J., Rubin B.Y.;
RT "Familial dysautonomia is caused by mutations of the IKAP gene.";
RL Am. J. Hum. Genet. 68:753-758(2001).
CC -!- FUNCTION: MAY ACT AS A SCAFFOLD PROTEIN THAT MAY ASSEMBLE ACTIVE
CC IKK-MAP3K14 COMPLEXES (IKKA, IKKB AND MAP3K14/NIK).
CC -!- SUBUNIT: INTERACTS PREFERENTIALLY WITH MAP3K14/NIK FOLLOWED BY
CC IKK-ALPHA AND IKK-BETA.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- PTM: PHOSPHORYLATED.
CC -!- DISEASE: DEFECTS IN IKKAP ARE THE CAUSE OF FAMILIAL DYSAUTONOMIA
CC (FD); ALSO KNOWN AS RILEY-DAY SYNDROME OR HEREDITARY SENSORY AND
CC AUTONOMIC NEUROPATHY TYPE III. THIS AUTOSOMAL RECESSIVE DISORDER
CC IS DUE TO THE POOR DEVELOPMENT AND SURVIVAL, AND PROGRESSIVE
CC DEGENERATION OF THE SENSORY, SYMPATHETIC AND PARASYMPATHETIC
CC NEURONS. FD INDIVIDUALS ARE AFFECTED WITH A VARIETY OF SYMPTOMS
CC SUCH AS DECREASED SENSITIVITY TO PAIN AND TEMPERATURE,
CC CARDIOVASCULAR INSTABILITY, RECURRENT PNEUMONIAS, VOMITING CRISES,
CC AND GASTROINTESTINAL DYSFUNCTION. IT IS PRIMARILY CONFINED TO
CC INDIVIDUALS OF ASHKENAZI JEWISH DESCENT, WITH AN INCIDENCE OF
CC 1/3600 LIVE BIRTHS.
CC -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 1286.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF044195; AAC64258.1; -
 DR EMBL; AF153419; AAG43369.1; -
 DR EMBL; AL049945; CAB43219.1; ALT_FRAME.
 DR Genew; HGNC:5959; IAKKAP.
 DR MIM; 603722; -
 DR MIM; 223900; -
 DR GO; GO:0008607; F:phosphorylase kinase, regulator activity; TAS.
 DR GO; GO:0004871; F:signal transducer activity; TAS.
 DR GO; GO:0006955; P:immune response; TAS.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
 DR GO; GO:0006461; P:protein complex assembly; TAS.
 DR InterPro; IPR006849; IKI3.
 DR Pfam; PF04762; IKI3; 1.
 KW Phosphorylation; Disease mutation.
 FT VARIANT 696 696 R -> P (IN FD; MILD PHENOTYPE;
 FT PHOSPHORYLATION IS REDUCED).
 FT /FTID=VAR_011327.
 FT CONFLICT 304 304 R -> W (IN REF. 2).
 FT CONFLICT 312 312 K -> E (IN REF. 2).
 FT CONFLICT 754 754 P -> L (IN REF. 2).
 FT CONFLICT 961 961 C -> G (IN REF. 3).
 FT CONFLICT 1072 1072 S -> C (IN REF. 1).
 FT CONFLICT 1158 1158 P -> L (IN REF. 3).
 FT CONFLICT 1320 1320 I -> V (IN REF. 3).
 SQ SEQUENCE 1332 AA; 3FD65FAB554DA923 CRC64;

Query Match 65.4%; Score 34; DB 1; Length 1332;
 Best Local Similarity 60.0%; Pred. No. 78;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 MKAFVFEED 10
 Db. 1240 LKVLFLFEED 1249

RESULT 4

PHOR_ECOLI STANDARD; PRT; 431 AA.
 AC P08400;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phosphate regulon sensor protein phor (EC 2.7.3.-).
 GN PHOR OR NWPB OR B0400.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=87169739; PubMed=3550103;
 RA Makino K., Shinagawa H., Anemura M., Nakata A.;
 RT "Nucleotide sequence of the phoR gene, a regulatory gene for the
 RT phosphate regulon of Escherichia coli.";
 RL J. Mol. Biol. 192:549-556(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,

RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H.,
 RA Lin D., Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SUBCELLULAR LOCATION.
 RX MEDLINE=90251245; PubMed=2187152;
 RA Yamada M., Makino K., Shinagawa H., Nakata A.;
 RT "Regulation of the phosphate regulon of Escherichia coli: properties
 RT of phoR deletion mutants and subcellular localization of PhoR
 RT protein.";
 RL Mol. Gen. Genet. 220:366-372(1990).
 RN [5]
 RP TOPOLOGY.
 RX MEDLINE=93302503; PubMed=8391104;
 RA Scholten M., Tomassen J.;
 RT "Topology of the PhoR protein of Escherichia coli and functional
 RT analysis of internal deletion mutants.";
 RL Mol. Microbiol. 8:269-275(1993).
 CC -!- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM PHOR/PHOB
 CC INVOLVED IN THE PHOSPHATE REGULON GENES EXPRESSION. PHOR MAY
 CC FUNCTION AS A MEMBRANE-ASSOCIATED PROTEIN KINASE THAT
 CC PHOSPHORYLATES PHOB IN RESPONSE TO ENVIRONMENTAL SIGNALS.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -!- SIMILARITY: Contains 1 histidine kinase domain.
 CC -!- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X04704; CAA28409.1; -
 DR EMBL; AE000146; AAC73503.1; -
 DR EMBL; U73857; AAB18124.1; -
 DR PIR; A25557; RGCFR.
 DR HSP; P02933; IJOY.
 DR EcoGene; EG10733; phoR.
 DR InterPro; IPR003594; ATPbind_ATPase.
 DR InterPro; IPR004358; Bact_sens_pr_C.
 DR InterPro; IPR003661; His_kinase.
 DR InterPro; IPR005467; His_kinase.
 DR InterPro; IPR000014; PAS_domain.
 DR Pfam; PF02518; HATPase_C; 1.
 DR Pfam; PF00512; HlsKA; 1.
 DR Pfam; PF00989; PAS; 1.
 DR PRINTS; PR00344; BCTRLSENSOR.
 DR SMART; SM00387; HATPase_C; 1.
 DR SMART; SM00388; HlsKA; 1.
 DR SMART; SM00091; PAS; 1.
 DR TIGRfams; TIGR00229; sensory_box; 1.
 DR PROSITE; PS50109; HIS_KIN; 1.
 DR PROSITE; PS50112; PAS; 1.
 KW Sensory transduction; Transferase; Kinase; Phosphorylation;
 KW Transmembrane; Inner membrane; Phosphate transport; Complete proteome.
 FT DOMAIN 1 9 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 10 28 PROBABLE.
 FT DOMAIN 29 32 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 33 51 PROBABLE.
 FT DOMAIN 52 431 CYTOPLASMIC (PROBABLE).
 FT DOMAIN 96 172 PAS.
 FT DOMAIN 210 425 HISTIDINE KINASE.
 FT MOD_RES 213 213 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 SQ SEQUENCE 431 AA; 49629 MW; 33883582AF4B883C CRC64;

Query Match 63.5%; Score 33; DB 1; Length 431;
 Best Local Similarity 66.7%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 KOAFVFEED 10
 |||||

Db 298 KQTFEID 306

RESULT 5

PHOR_SHIDY STANDARD; PRT; 431 AA.

AC P45609;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Phosphate regulon sensor protein phor (EC 2.7.3.-).

GN PHOR.

OS Shigella dysenteriae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Shigella.

OX NCBI_TaxID=622;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90078103; PubMed=2556368;

RA Lee T.Y., Makino K., Shinagawa H., Amemura M., Nakata A.;

RT "Phosphate regulon in members of the family Enterobacteriaceae: comparison of the phoB-phoR operons of Escherichia coli, Shigella dysenteriae, and Klebsiella pneumoniae.";

RL J. Bacteriol. 171:6593-6599(1989).

CC -!- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM PHOR/PHOB INVOLVED IN THE PHOSPHATE REGULATED GENES EXPRESSION. PHOR MAY FUNCTION AS A MEMBRANE-ASSOCIATED PROTEIN KINASE THAT PHOSPHORYLATES PHOB IN RESPONSE TO ENVIRONMENTAL SIGNALS.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.

CC -!- SIMILARITY: Contains 1 histidine kinase domain.

CC -!- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M31793; AAA26536.1; -

DR PIR; B44753; B44753.

DR HSSP; P02933; IJOY.

DR InterPro; IPR003594; ATPbind_ATPase.

DR InterPro; IPR003561; His_kinase.

DR InterPro; IPR005467; His_kinase.

DR InterPro; IPR000014; PAS_domain.

DR Pfam; PF02518; HATPase_c; 1.

DR Pfam; PF00512; Hiska; 1.

DR Pfam; PF00989; PAS; 1.

DR SMART; SM00387; HATPase_c; 1.

DR SMART; SM00388; Hiska; 1.

DR SMART; SM00091; PAS; 1.

DR TIGRFAMs; TIGR00229; sensory_box; 1.

DR PROSITE; PS50109; HIS_KIN; 1.

DR PROSITE; PS50112; PAS; 1.

KW Sensory transduction; Transferase; Kinase; Phosphorylation;

KW Transmembrane; Inner membrane; Phosphate transport.

FT DOMAIN 1 13 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 14 34 POTENTIAL.

FT DOMAIN 35 38 PERIPLASMIC (POTENTIAL).

FT TRANSMEM 39 59 POTENTIAL.

FT DOMAIN 60 431 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 96 167 PAS.

FT DOMAIN 210 425 HISTIDINE KINASE.

FT MOD_RES 213 213 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

SQ SEQUENCE 431 AA; 49702 MW; 0D0D84FC268253E0 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 431;

Best Local Similarity 66.7%; Pred. No. 42;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 KQAFVFEID 10.

Db 298 KQTFEID 306

RESULT 6

THIP_HAEIN STANDARD; PRT; 538 AA.

AC P44985;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Thiamine transport system permease protein thip.

GN THIP OR H11020.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Haemophilus.

OX NCBI_TaxID=727;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=RD / KW20 / ATCC 51907;

RA MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,

RA Karlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,

RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,

RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,

RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,

RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

RA Venter J.C.;

RT "Whole-genome random sequencing and assembly of Haemophilus influenzae

RD.";

RL Science 269:496-512(1995).

CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM

CC TEPA-THIPO FOR THIAMINE AND TPP. PROBABLY RESPONSIBLE FOR THE

CC TRANSLOCATION OF THE SUBSTRATE ACROSS THE MEMBRANE (BY

CC SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane

CC (Potential).

CC -!- SIMILARITY: BELONGS TO THE BINDING-PROTEIN-DEPENDENT TRANSPORT

CC SYSTEM PERMEASE FAMILY. CYSTW SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U32782; AAC22679.1; -

CC PIR; D64164; D64164.

CC TIGR; H11020; -

DR InterPro; IPR000515; BPD_transp.

DR InterPro; IPR005947; Thip_ABC_transp.

DR Pfam; PF00528; BPD_transp; 1.

DR TIGRFAMs; TIGR01253; thip; 1.

DR PROSITE; PS00402; BPD_TRANSP_INN_MEMBR; FALSE_NEG.

DR Transmembrane; Inner membrane; Transport; Complete proteome.

FT TRANSMEM 19 59 POTENTIAL.

FT TRANSMEM 58 78 POTENTIAL.

FT TRANSMEM 97 117 POTENTIAL.

FT TRANSMEM 141 161 POTENTIAL.

FT TRANSMEM 202 222 POTENTIAL.

FT TRANSMEM 242 262 POTENTIAL.

FT TRANSMEM 293 313 POTENTIAL.

FT TRANSMEM 337 357 POTENTIAL.

FT TRANSMEM 376 396 POTENTIAL.

FT TRANSMEM 406 426 POTENTIAL.

FT TRANSMEM 466 486 POTENTIAL.

FT TRANSMEM 509 529 POTENTIAL.

SQ SEQUENCE 538 AA; 60851 MW; D7B31D2A15BCD6E1 CRC64;

```

Query Match      63.5%; Score 33; DB 1; Length 538;
Best Local Similarity 75.0%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 QAFVFEFD 10
      II :IIII
DB      234 QAILFEFD 241

RESULT 7
GPH_SYNP7
ID GPH_SYNP7 STANDARD; PRT; 212 AA.
AC Q55039;
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoglycolate phosphatase (EC 3.1.3.18) (PGP).
GN CBBZ.
OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
ON NCBI_TaxID=11140;
RX SEQUENCE FROM N.A.
RP MEDLINE=96257262; PubMed=8654968;
RA Liu Y., Tsinoiremas N.F.;
RT "An unusual gene arrangement for the putative chromosome replication
RT origin and circadian expression of dnaN in Synecococcus sp. strain
RT PCC 7942.";
RL Gene 172:105-109(1996).
CC -!- CATALYTIC ACTIVITY: 2-phosphoglycolate + H(2)O -> glycolate +
CC phosphate.
CC -!- SIMILARITY: BELONGS TO THE CBBY/CBBZ/GPH/YIEH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U33322; AAA75108.1;
DR HAWAP; MF_00495; atypical; 1.
DR InterPro; IPR005833; Hignase/hydrase.
DR InterPro; IPR005834; Hydrolase.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00413; HADHALOGNASE.
KW Carbohydrate metabolism; Calvin cycle; Hydrolase.
SQ SEQUENCE 212 AA; 23572 MW; 8054F39165DECECF CRC64;

Query Match      61.5%; Score 32; DB 1; Length 212;
Best Local Similarity 62.5%; Pred. No. 34;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      3 QAFVFEFD 10
      II :IIII
DB      2 QAILFEFD 9

RESULT 8
Y414_METJA
ID Y414_METJA STANDARD; PRT; 395 AA.
AC Q57857;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0414.
GN MJ0414.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]

```

```

SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67493; AAB98403.1;
DR PIR; F64351; F64351.
DR TIGR; MJ0414;
DR InterPro; IPR001072; Y414.
DR Pfam; PF02003; DUF135; 1.
DR PRINTS; PR01048; Y414FAMILY.
DR ProDom; PD014265; Y414; 1.
DR TIGRFAMs; TIGR01209; TIGR01209; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 395 AA; 46417 MW; 47AD62037D44D33B CRC64;

Query Match      61.5%; Score 32; DB 1; Length 395;
Best Local Similarity 60.0%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 MKQAFVFEFD 10
      :IIIIII
DB      262 LKSATFFFD 271

RESULT 9
ECSB_BACSU
ID ECSB_BACSU STANDARD; PRT; 408 AA.
AC P55340;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Protein ecsb.
GN ECSB OR PRST.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96146051; PubMed=8581172;
RA Leskela S., Kontinen V.P., Sarvas M.;
RT "Molecular analysis of an operon in Bacillus subtilis encoding a
RT novel ABC transporter with a role in exoprotein production,
RT sporulation and competence.";
RL Microbiology 142:71-77(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98240224; PubMed=9579061;
RA Noback M.A., Holsappel S., Kiewiet R., Terpstra P., Wambutt R.,
RA Wedler H., Venema G., Bron S.;
RT "The 172 kb prkA-addAB region from 83 degrees to 97 degrees of the
RT Bacillus subtilis chromosome contains several dysfunctional genes,
RT the glyB marker, many genes encoding transporter proteins, and the

```

```

RT ubiquitous hit gene.
RL Microbiology 144:859-875(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Broussier L., Bruns A., Braun M., Brignell S.C., Bron S.,
RA Bouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dunthorn J., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Ertington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Chim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivoita C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Takakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weltzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -!- FUNCTION: Presumed to form part of an ABC-transporter, it may form
CC a transport channel.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X87807; CAA61075.1;
CC EMBL: Y14077; CAA74408.1;
CC EMBL: 299109; CAB12845.1;
CC PIR: G69619; G69619.
CC Subtilist; Bg11517; ecsB.
CC Competence; Sporulation; Transport; Transmembrane; Complete proteome.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 53 73 POTENTIAL.
FT TRANSMEM 111 131 POTENTIAL.
FT TRANSMEM 134 154 POTENTIAL.
FT TRANSMEM 180 200 POTENTIAL.
FT TRANSMEM 284 304 POTENTIAL.
FT TRANSMEM 308 328 POTENTIAL.
FT TRANSMEM 351 371 POTENTIAL.
FT TRANSMEM 374 394 POTENTIAL.
SQ SEQUENCE 408 AA; 47299 MW; 344628E5A3E7DAAE CRC64;

Query Match 61.5%; Score 32; DB 1; Length 408;
Best Local Similarity 55.6%; Pred. No. 62;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MKQAFVFEF 9
:||||:
Db 99 LKQAFVYSY 107

RESULT 11
LP1A_DROME
ID LP1A_DROME STANDARD: PRT: 789 AA.
AC P11995; Q9VVM4;
DT 01-OCT-1989 (Rel. 12, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Larval serum protein 1 alpha chain precursor (Hexamerin 1 alpha).
GN LSP1-ALPHA OR LSP1-A OR CG2559.

Query Match 61.5%; Score 32; DB 1; Length 744;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 MKQAFVFEFD 10
:||||:
Db 481 VKKAFSFD 490

RESULT 10
CATA_EMENI
ID CATA_EMENI STANDARD: PRT: 744 AA.
AC P55305;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Catalase A (EC 1.11.1.6) (Spore-specific catalase).
GN CATA;
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=162425;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FGSC 26;
RX MEDLINE=96171518; PubMed=8598056;
RA Navarro R.E., Stringer M.A., Hansberg W., Timberlake W.E., Aguirre J.;
RA "cata", a new Aspergillus nidulans gene encoding a developmentally
RA regulated catalase.
RL Curr. Genet. 29:352-359(1996).
CC -!- FUNCTION: Occurs in almost all aerobically respiring organisms and
CC serves to protect cells from the toxic effects of hydrogen
CC peroxide.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP.
CC -!- DEVELOPMENTAL STAGE: SPORULATION-SPECIFIC.
CC -!- SIMILARITY: Belongs to the catalase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U37803; AAC49254.1;
CC PIR: S68115; S68115.
CC HSP; P21179; ICF9.
CC InterPro: IPR002226; Catalase.
CC Pfam: PF00199; catalase; 1.
CC PRINTS; PR00067; CATALASE.
CC ProDom: PD000510; Catalase; 1.
CC PROSITE; PS00437; CATALASE_1; 1.
CC PROSITE; PS00436; CATALASE_2; 1.
CC Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Sporulation.
FT ACT_SITE 93 93 BY SIMILARITY.
FT ACT_SITE 166 166 BY SIMILARITY.
FT METAL 380 380 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 744 AA; 84063 MW; 4B4D0B7C8EC8DEFA CRC64;

Query Match 61.5%; Score 32; DB 1; Length 744;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 MKQAFVFEFD 10
:||||:
Db 481 VKKAFSFD 490

```

OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.C., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.B.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Poirard J., Puri V., Reese M.G.,
 RA Reinartz K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang J., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE OF 1-84 FROM N.A.
 RX MEDLINE=87060914; PubMed=3097321;
 RA Delaney S.J., Smith D.F., McClelland A., Sunkel C., Glover D.M.;
 RA "Sequence conservation around the 5' ends of the larval serum protein
 RT 1 genes of Drosophila melanogaster";
 RL J. Mol. Biol. 189:1-11(1986).
 RN [3]
 RP SEQUENCE OF 1-52 FROM N.A.
 RX Jovett T.;
 RA "The regulatory domain of a larval serum protein gene in Drosophila
 RL melanogaster";
 RL EMBO J. 4:3789-3795(1985).
 CC -!- FUNCTION: LARVAL STORAGE PROTEIN (LSP) WHICH MAY SERVE AS A STORE
 CC OF AMINO ACIDS FOR SYNTHESIS OF ADULT PROTEINS (BY SIMILARITY).
 CC -!- SUBUNIT: HETEROHEXAMER, COMPOSED OF THREE SUBUNITS, ALPHA, BETA
 CC AND GAMMA.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- TISSUE SPECIFICITY: LARVAL HEMOLYMPH.
 CC -!- SIMILARITY: TO ARYLPHORINS AND TO ARTHROPOD HEMOCYANINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entitles requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE003489; AAF48168.1; -
 DR EMBL; X03872; CAA27506.1; -
 DR EMBL; X03368; CAA27066.1; -
 DR PIR; A27144; A27144.
 DR HSP; P04253; 1LLA.
 DR FlyBase; FBgn0002562; Lsp1-alpha.
 DR GO; GO:0005616; C:larval serum protein complex; IDA.
 DR InterPro; IPR000896; Hemocyanin.
 DR InterPro; IPR005203; hemocyanin_C.
 DR InterPro; IPR005204; hemocyanin_N.
 DR Pfam; PF00372; hemocyanin_1.
 DR Pfam; PF03723; hemocyanin_C; 1.
 DR Pfam; PF03722; hemocyanin_N; 1.
 DR PRINTS; PS00187; HAEMOCYANIN.
 DR PROSITE; PS00209; HEMOCYANIN_1; FALSE_NEG.
 DR PROSITE; PS00210; HEMOCYANIN_2; 1.
 KW Signal; Hemolymph; Storage protein; Glycoprotein; Multigene family.
 FT SIGNAL 1 16 POTENTIAL.
 FT CHAIN 17 789 LARVAL SERUM PROTEIN 1 ALPHA CHAIN.
 SQ SEQUENCE 789 AA; 95883 MW; 47F3F21B05D53795 CRC64;
 Query Match 61.5%; Score 32; DB 1; Length 789;
 Best Local Similarity 75.0%; Pred No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MKQAFVFE 8
 Db 34 MKQKFLFE 41
 RESULT 12
 LP1B_DROME STANDARD; PRT; 789 AA.
 ID PL1B_DROME
 AC PL1996; Q9VPV2;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Larval serum protein 1 beta chain (sequence update)
 GN LSP1-BETA OR LSP1-B OR CG4178.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Canton-S;
 RX MEDLINE=97274657; PubMed=9128742;
 RA Massey H.C. Jr., Kejzlarova-Lepesant J., Willis R.L.,
 RA Castlberry A.B., Benes H.;
 RA "The Drosophila Lsp-1 beta gene. A structural and phylogenetic
 RT analysis";
 RL Eur. J. Biochem. 245:199-207(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.C., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Morkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacht J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*,"
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE OF 1-100 FROM N.A.
RX MEDLINE=87060914; PubMed=3097321;
RA Delaney S.J., Smith D.F., McClelland A., Sunkel C., Glover D.M.;
RT "Sequence conservation around the 5' ends of the larval serum protein
RT 1 genes of *Drosophila melanogaster*,"
RL J. Mol. Biol. 189:1-11(1986).
CC -1- FUNCTION: LARVAL STORAGE PROTEIN (LSP) WHICH MAY SERVE AS A STORE
CC OF AMINO ACIDS FOR SYNTHESIS OF ADULT PROTEINS (BY SIMILARITY).
CC -1- SUBUNIT: HETEROHEXAMER, COMPOSED OF THREE SUBUNITS, ALPHA, BETA
CC AND GAMMA.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: LARVAL HEMOLYMPH.
CC -1- SIMILARITY: TO ARYLPHORINS AND TO ARTHROPOD HEMOCYANINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U63556; AB58821.1;
CC EMBL: AE003588; AAF51434.1;
CC EMBL: X03873; CAA27507.1;
CC PIR: B27144; B27144.
CC HSP: P04253; I0XY.
CC FlyBase: FBgn0002563; Lspl-beta.
CC Go: GO:0005016; C:larval serum protein complex; IDA.
CC InterPro: IPR000896; Hemocyanin.
CC InterPro: IPR005203; hemocyanin.C.
CC InterPro: IPR005204; hemocyanin.N.
CC Pfam: PF00372; hemocyanin.1.
CC Pfam: PF03723; hemocyanin.C.
CC Pfam: PF03722; hemocyanin.N.
CC PROSITE: PS00209; HEMOCYANIN.1; FALSE_NEG.
CC PROSITE: PS00210; HEMOCYANIN.2; 1.
CC Signal: Hemolymph; Storage protein; Multigene family.
CC SIGNAL 1 16 POTENTIAL
CC CHAIN 17 789 LARVAL SERUM PROTEIN 1 BETA CHAIN.
CC CONFLICT 275 275 T -> N (IN REF. 1).
CC CONFLICT 725 725 S -> R (IN REF. 1).
CC SEQUENCE 789 AA; 95913 MW; AE12594515806A5B CRC64;

Query Match 61.5%; Score 32; DB 1; Length 789;
Best Local Similarity 75.08; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Caps 0;

QY 1 MKQAFVFE 8
DB 34 MKQKFLFE 41
RESULT 13
SYFB_BACSU STANDARD; PRT; 804 AA.
AC P17922: P94540;
DT 01-NOV-1990 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 41, Last annotation update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)
DE (Phenylalanyl-tRNA synthetase beta chain) (PHERS).
GN PHER.
OS *Bacillus subtilis*.
OC Bacteria; Firmicutes; Bacillales; ~~Bacillaceae~~; *Bacillus*.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=91175935; PubMed=2127701;
RT Brakhage A., Wozny M., Putzer H.;
RT "Structure and nucleotide sequence of the *Bacillus subtilis*
RT phenylalanyl-tRNA synthetase genes,"
RL Biochimie 72:725-734(1990).
RN [2]
RP ERRATUM.
RX MEDLINE=91234765; PubMed=1903307;
RA Brakhage A., Wozny M., Putzer H.;
RL Biochimie 73:127-127(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97124191; PubMed=8969504;
RA Wipat A., Carter N., Brignell C.S., Guy J.B., Piper K.,
RA Sanders J., Emmerson P.T., Harwood C.R.;
RA "The dnaB-pheA (256 degrees-240 degrees) region of the *Bacillus*
RT subtilis chromosome containing genes responsible for stress
RT responses, the utilization of plant cell walls and primary
RT metabolism,"
RL Microbiology 142:3067-3078(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriass R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haelech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinot S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche M., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RL *subtilis*."; NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Pm70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.,
 RT "Complete genomic sequence of *Pasteurella multocida* Pm70."; Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 RL CC -!- CATALYTIC ACTIVITY: ATP + N(2)-formyl-N(1)-(5-phospho-D-
 CC ribosyl)glycinamide + L-glutamine + H(2)O = ADP + phosphate + 2-
 CC (formamido)-N(1)-(5-phospho-D-riboseyl)acetamidine + L-glutamate.
 CC -!- PATHWAY: De novo purine biosynthesis; fourth step.
 CC -!- SUBUNIT: Monomer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE FGAMS
 CC FAMILY.
 CC -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 DR EMBL: X53057; CAA37225.1; -
 DR EMBL: 275208; CAA99564.1; -
 DR EMBL: 299118; CAB14823.1; -
 DR PIR: A69676; YFBSB.
 DR HSSP: P27002; 1PVS.
 DR Subtilist: BG10875; phet.
 DR HAMAP: MF_00283; 1.
 DR InterPro: IPR005146; B3_4.
 DR InterPro: IPR005147; B5.
 DR InterPro: IPR005121; Fdx-AnticB.
 DR InterPro: IPR004532; phet_Bact.
 DR InterPro: IPR002547; tRNA_bind.
 DR Pfam: PF03483; B3_4; 1.
 DR Pfam: PF03484; B5; 1.
 DR Pfam: PF03147; FDX-ACB; 1.
 DR Pfam: PF01588; tRNA_bind; 1.
 DR TIGRFAMS: TIGR00472; phet_bact; 1.
 DR PROSITE: PS50886; TRBD; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Metal-binding; Magnesium; RNA-binding; tRNA-binding;
 KW Complete proteome.
 FT DOMAIN 40 155 TRNA-BINDING.
 FT METAL 462 462 MAGNESIUM (BY SIMILARITY).
 FT METAL 468 468 MAGNESIUM (VIA CARBONYL OXYGEN)
 FT METAL 468 468 (BY SIMILARITY).
 FT METAL 471 471 MAGNESIUM (BY SIMILARITY).
 FT METAL 472 472 MAGNESIUM (BY SIMILARITY).
 FT CONFLICT 110 111 KL -> NV (IN REF. 1).
 SQ SEQUENCE 804 AA; 87945 MW; D237DEAE639E551 CRC64;
 Query Match 61.5%; Score 32; DB 1; Length 804;
 Best Local Similarity 50.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MKQAFVEEFD 10
 Db 683 IKETTVFELD 692
 RESULT 14
 PUR4_PASMU
 ID PUR4_PASMU STANDARD; PRT; 1297 AA.
 AC Q9CLW4;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phosphoribosylformylglycinamide synthase (EC 6.3.5.3) (FGAM
 DE synthase) (FGAMS) (formylglycinamide ribotide amidotransferase)
 DE (FGARAT) (formylglycinamide ribotide synthetase).
 GN PURL OR PM1085.

OS *Pasteurella multocida*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Pm70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.,
 RT "Complete genomic sequence of *Pasteurella multocida* Pm70."; Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 RL CC -!- CATALYTIC ACTIVITY: ATP + N(2)-formyl-N(1)-(5-phospho-D-
 CC ribosyl)glycinamide + L-glutamine + H(2)O = ADP + phosphate + 2-
 CC (formamido)-N(1)-(5-phospho-D-riboseyl)acetamidine + L-glutamate.
 CC -!- PATHWAY: De novo purine biosynthesis; fourth step.
 CC -!- SUBUNIT: Monomer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE FGAMS
 CC FAMILY.
 CC -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 DR EMBL: AE006149; AAK03169.1; -
 DR HAMAP: MF_00419; 1.
 DR InterPro: IPR000728; AIRS-related.
 DR Pfam: PF00586; AIRS; 1.
 DR Pfam: PF02769; AIRS_C; 2.
 KW Purine biosynthesis; Ligase; ATP-binding; Glutamine amidotransferase;
 KW Complete proteome.
 FT NP_BIND 307 318 ATP (POTENTIAL).
 FT ACT_SITE 1137 1137 GATASE (BY SIMILARITY).
 SQ SEQUENCE 1297 AA; 143042 MW; 9362706FEFE34CB4 CRC64;
 Query Match 61.5%; Score 32; DB 1; Length 1297;
 Best Local Similarity 85.7%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 AFVEEFD 10
 Db 111 AFYFEED 117
 RESULT 15
 YG21_YEAST
 ID YG21_YEAST STANDARD; PRT; 283 AA.
 AC P53234;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Hypothetical 32.2 kDa protein in SCM4-MU1 intergenic region.
 GN YGR053C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Entian K.D., Rose M., Koetter P., Roehmer A., Sehrsam I.,
 RA Hempel S.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; 272838; CAA97053.1; -
DR PIR; S64347; S64347.
DR SGD; S0003285; YGR053C.
KW Hypothetical protein.
SQ SEQUENCE 283 AA; 32200 MW; 750BF7929396C590 CRC64;

Query Match 59.6%; Score 31; DB 1; Length 283;
Best Local Similarity 50.0%; Pred. No. 69;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 MKQAFVFEFD 10

Db 243 IKGTFFIEFN 252

Search completed: July 24, 2003, 20:06:32
Job time : 4.75325 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:01:59 ; Search time 7.5974 seconds
(without alignments)
339.659 Million cell updates/sec

Title: US-09-546-136-1

Perfect score: 52

Sequence: 1 MKQAFVFEFD 10

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	94.2	144	2 052975	052975 clostridium
2	41	78.8	452	10 Q9L173	Q9L173 arabidopsis
3	37	71.2	142	2 050600	050600 clostridium
4	36	69.2	962	12 Q9DUT3	Q9DUT3 pelargonium
5	35	67.3	761	10 Q94J14	Q94J14 oryza sativ
6	35	67.3	1754	5 Q8I2Q1	Q8I2Q1 plasmodium
7	34	65.4	285	17 Q97VC7	Q97VC7 sulfolobus
8	34	65.4	374	2 Q9F8S3	Q9F8S3 thermus fil
9	34	65.4	1332	4 Q8N516	Q8N516 homo sapien
10	34	65.4	1333	6 Q8WND5	Q8WND5 oryctolagus
11	33	63.5	71	12 Q91FQ1	Q91FQ1 chilo iride
12	33	63.5	420	10 Q93YPL	Q93YPL arabidopsis
13	33	63.5	431	16 Q8XE99	Q8XE99 escherichia
14	33	63.5	431	16 Q8FKD0	Q8FKD0 escherichia
15	33	63.5	436	2 Q9RHV5	Q9RHV5 streptococc
16	33	63.5	436	16 Q8DS90	Q8DS90 streptococc

ALIGNMENTS

RESULT 1

O52975 PRELIMINARY: PRT: 144 AA.
AC O52975;
DT 01-JUN-1998 (TremBLrel. 06, Created)
DT 01-JUN-1998 (TremBLrel. 06, Last sequence update)
DT 01-NOV-1998 (TremBLrel. 08, Last annotation update)
DE ORF-XI.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-type E Iwanai.
RX MEDLINE=98126542; PubMed=9465394;
RA Kubota T., Yonekura N., Hariya Y., Isogai E., Isogai H., Amano K.,
RA Fujii N.;
RT "Gene arrangement in the upstream region of Clostridium botulinum type E and Clostridium butyricum BL6340 progenitor toxin genes is different from that of other types.";
RL FEMS Microbiol. Lett. 158:215-221(1998).
DR EMBL: D88418; BAA24880.1;
SQ SEQUENCE 144 AA: 16767 MW: 3447DB7F67670BC CRC64;

Query Match 94.2%; Score 49; DB 2; Length 144;

Best Local Similarity 90.0%; Pred. No. 0.027;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKQAFVFEFD 10

:|||||

3 LKQAFVFEFD 12

RESULT 2

Q9L173 PRELIMINARY: PRT: 452 AA.

ID Q9L173

AC Q9L173

DT 01-OCT-2000 (TremBLrel. 15, Created)

```

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Chloroplast nucleoid DNA binding protein-like, nucleolin-like
DE protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Columbia;
RC Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=Columbia;
RC MEDLINE=20363099; PubMed=10907853;
RX Nakamura Y.;
RA "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT TAC and BAC clones";
RL DNA Res. 7:217-221(2000).
DR EMBL; AP001313; BAB03090.1;
DR HSP; P00797; 2REN.
DR InterPro; IPR001461; AspproteaseA1.
DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
SQ SEQUENCE 452 AA; 49374 MW; DDCF09D45B566410 CRC64;

Query Match 78.8%; Score 41; DB 10; Length 452;
Best Local Similarity 70.0%; Pred. No. 3.7;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Caps 0;

OY 1 MKQAFVFEED 10
DB 427 MQQGLFEEF 436

RESULT 3
O50600
ID O50600 PRELIMINARY; PRT; 142 AA.
AC O50600;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE ORF-XI.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=type F Langeland;
RX MEDLINE=98126542; PubMed=9465394;
RA Kubota T., Yonekura N., Hariya Y., Isogai E., Isogai H., Amano K.,
RA Fujii N.;
RT "Gene arrangement in the upstream region of Clostridium botulinum type
RT E and Clostridium butyricum BL6340 progenitor toxin genes is different
RT from that of other types.";
RL FEMS Microbiol. Lett. 158:215-221(1998).
DR EMBL; AB004779; BAA24890.1;
SQ SEQUENCE 142 AA; 16539 MW; 8C0C682EB824122F CRC64;

Query Match 71.2%; Score 37; DB 2; Length 142;
Best Local Similarity 70.0%; Pred. No. 8.1;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Caps 0;

OY 1 MKQAFVFEED 10
DB 1 MKQTFSFNFD 10

```

```

RESULT 4
O9DUT3
ID O9DUT3 PRELIMINARY; PRT; 962 AA.
AC O9DUT3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Genomic RNAI for Ia protein, (ORF1).
OS Pelargonium zonate spot virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae.
OX NCBI_TaxID=116056;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=tomato;
RC Gallitelli D., Finetti-Sialer M.M.;
RA "Complete nucleotide sequence of Pelargonium zonate spot virus and
RT it's relationship with the bromoviridae family.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ272327; CAC08526.1;
DR InterPro; IPR000606; Viral_helicase1.
DR InterPro; IPR002588; V_methyltransf.
DR Pfam; PF01443; Viral_helicase1; 1.
DR Pfam; PF01660; Vmethyltransf; 1.
SQ SEQUENCE 962 AA; 108489 MW; B545F0F84858A891 CRC64;

Query Match 69.2%; Score 36; DB 12; Length 962;
Best Local Similarity 70.0%; Pred. No. 83;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Caps 0;

OY 1 MKQAFVFEED 10
DB 706 IKRAFVFESD 715

RESULT 5
O94J14
ID O94J14 PRELIMINARY; PRT; 761 AA.
AC O94J14;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE P0638D12.5 protein (OSJNB0021A09.19 protein).
GN P0638D12.5 OR OSJNB0021A09.19.
OS Oryza sativa (Rice), and
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530, 39947;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Nipponbare;
RC Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone: P0638D12.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Nipponbare;
RC Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
RT clone: OSJNB0021A09.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002972; BAB55497.1;
DR EMBL; AP003218; BAB89467.1;
DR Gramene; O94J14;
DR InterPro; IPR001711; PL_PLC_Y.
DR PROSITE; PSS0008; PIPLC_Y_DOMAIN; 1.
SQ SEQUENCE 761 AA; 85474 MW; 992666EDB3EC9B9DF CRC64;

Query Match 67.3%; Score 35; DB 10; Length 761;

```

```

Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KQAFVFEF 9
DB 563 KQALVFEF 570
   ||| ||||
   ||| ||||

RESULT 6
Q81201 PRELIMINARY; PRT; 1754 AA.
AC Q81201:
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF11265W.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowden S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes K.D., Lark N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinovitch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."
RL Nature 419:527-531(2002).
DR EMBL; AL929358; CAD51939.1;
KW Hypothetical protein.
SQ SEQUENCE 1754 AA; 208589 MW; CB074AE4C3D99DE CRC64;

Query Match 67.3%; Score 35; DB 5; Length 1754;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KQAFVFEF 10
DB 335 KDNFIFEFD 343
   | | ||||
   | | ||||

RESULT 7
Q97VC7 PRELIMINARY; PRT; 285 AA.
ID Q97VC7:
AC Q97VC7:
DT 01-OCT-2001 (TRENBLrel. 18, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
DE Regucalcin homolog.
GN SS02705.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Awayez M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,

```

```

RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duquet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AE006864; AAK42817.1;
DR InterPro; IPR005511; SMP-30.
DR Pfam; PF03758; SMP-30; 1.
KW Complete proteome.
SQ SEQUENCE 285 AA; 32194 MW; 3FA045545790DA5B CRC64;

Query Match 65.4%; Score 34; DB 17; Length 285;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKQAFVFEFD 10
DB 165 VRKVFVFEFD 174
   :: |||::|
   :: |||::|

RESULT 8
Q9F8S3 PRELIMINARY; PRT; 374 AA.
ID Q9F8S3:
AC Q9F8S3:
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE TflI restriction endonuclease.
GN TFIIR.
OS Thermus filiformis.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=276;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu S.-Y., Hsieh P.-C.;
RT "Method for cloning and producing the TflI restriction endonuclease in
RT E. coli."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF239924; AAG3969.1;
DR InterPro; IPR002294; D12N6_mtfrase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF02086; Methyltransfd12; 1.
DR PRINTS; PR00505; D12N6MTFRASE.
DR PROSITE; PS00092; N6_MTASE; 1.
KW Endonuclease.
SQ SEQUENCE 374 AA; 43895 MW; 3480B426FC1DD27C CRC64;

Query Match 65.4%; Score 34; DB 2; Length 374;
Best Local Similarity 50.0%; Pred. No. 86;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKQAFVFEFD 10
DB 326 MKKGFIVTFD 335
   ||: ||: ||
   ||: ||: ||

RESULT 9
Q8N516 PRELIMINARY; PRT; 1332 AA.
ID Q8N516:
AC Q8N516:
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase
DE complex-associated protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033094; AAH33094.1;
KW Kinase.
SQ SEQUENCE 1332 AA; 150223 MW; 55509840671D60CC CRC64;

Query Match 65.4%; Score 34; DB 4; Length 1332;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 MKOAFVFEED 10
   :| 1:||||
Db 1240 LKVLFLFEED 1249

RESULT 10
QBWNDS PRELIMINARY; PRT; 1333 AA.
AC QBWNDS;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE IKAP.
GN IKKAP.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21579790; PubMed=11722848;
RA Coli R., Anderson S.L., Volpi S.A., Rubin B.Y.;
RT "Genomic organization and chromosomal localization of the mouse IKKAP
   gene.";
RL Gene 279:81-89(2001).
DR EMBL; AF388202; AAL40927.1;
DR InterPro; IPR006849; IK13.
DR Pfam; PF04762; IK13; 1.
SQ SEQUENCE 1333 AA; 150794 MW; 00FA2FEE7046CC5F CRC64;

Query Match 65.4%; Score 34; DB 6; Length 1333;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 MKOAFVFEED 10
   :| 1:||||
Db 1241 LKVLFLFEED 1250

RESULT 11
Q91F01 PRELIMINARY; PRT; 71 AA.
AC Q91F01;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE 272L.
OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=10488;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H., Darai G., Fluegel R.M.;
RT "DNA analysis of insect iridescent virus 6: evidence for circular
   permutation and terminal redundancy.";
RL J. Virol. 49:609-614(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86174607; PubMed=3959991;
RA Lorbacher de Ruiz H., Gelderblom H., Hofmann W., Darai G.;
RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis
   in mice.";
RN [12]

```

```

RL Med. Microbiol. Immunol. 175:43-53(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=87321126; PubMed=2820141;
RA Schnitzler P., Soltan J.B., Fischer M., Reisner H., Scholz J.,
RA Delius H., Darai G.;
RT "Molecular cloning and physical mapping of the genome of insect
   iridescent virus type 6: further evidence for circular permutation of
   the viral genome.";
RL Virology 160:66-74(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89073752; PubMed=3201750;
RA Fischer M., Schnitzler P., Delius H., Darai G.;
RT "Identification and characterization of the repetitive DNA element in
   the genome of insect iridescent virus type 6.";
RL Virology 167:485-496(1988).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=92196996; PubMed=1549908;
RA Handermann M., Schnitzler P., Rosen-Wolff A., Raab K., Sonntag K.C.,
RA Darai G.;
RT "Identification and mapping of origins of DNA replication within the
   DNA sequences of the genome of insect iridescent virus type 6.";
RL Virus Genes 6:19-32(1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=93118242; PubMed=1475907;
RA Sonntag K.C., Darai G.;
RT "Characterization of the third origin of DNA replication of the genome
   of insect iridescent virus type 6.";
RL Virus Genes 6:333-342(1992).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=93260401; PubMed=8492091;
RA Stohwasser R., Raab K., Schnitzler P., Janssen W., Darai G.;
RT "Identification of the gene encoding the major capsid protein of
   insect iridescent virus type 6 by polymerase chain reaction.";
RL J. Gen. Virol. 74:873-879(1993).
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=94167241; PubMed=8121799;
RA Schnitzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,
RA Delius H., Darai G.;
RT "Identification of genes encoding zinc finger proteins, non-histone
   chromosomal HMG protein homologue, and a putative GTP phosphohydrolase
   in the genome of Chilo iridescent virus.";
RL Nucleic Acids Res. 22:158-166(1994).
RN [9]
RP SEQUENCE FROM N.A.
RX MEDLINE=94353641; PubMed=8073636;
RA Sonntag K.C., Schnitzler P., Koonin E.V., Darai G.;
RT "Chilo iridescent virus encodes a putative helicase belonging to a
   distinct family within the 'DEAD/H' superfamily: implications for the
   evolution of large DNA viruses.";
RL Virus Genes 8:151-158(1994).
RN [10]
RP SEQUENCE FROM N.A.
RX MEDLINE=95213160; PubMed=7698884;
RA Sonntag K.C., Schnitzler P., Janssen W., Darai G.;
RT "Identification of the primary structure and the coding capacity of
   the genome of insect iridescent virus type 6 between the genome
   coordinates 0.310 and 0.347 (7990 bp).";
RL Intervirology 37:287-297(1994).
RN [11]
RP SEQUENCE FROM N.A.
RX MEDLINE=94292906; PubMed=8021587;
RA Schnitzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,
RA Koonin E.V., Darai G.;
RT "Insect iridescent virus type 6 encodes a polypeptide related to the
   largest subunit of eukaryotic RNA polymerase II.";
RL J. Gen. Virol. 75:1557-1567(1994).
RN [12]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=98141693; PubMed=9482589;
RA Bahr U., Tidona C.A., Darai G.;
RT "The DNA sequence of Chilo iridescent virus between the genome
RT coordinates 0.101 and 0.391; similarities in coding strategy between
RT insect and vertebrate iridoviruses.";
RL Virus Genes 15:235-245(1997).
RN [13]
RP SEQUENCE FROM N.A.
RX MEDLINE=99125223; PubMed=9926400;
RA Muller K., Tidona C.A., Bahr U., Darai G.;
RT "Identification of a thymidylate synthase gene within the genome of
RT Chilo iridescent virus.";
RL Virus Genes 17:243-258(1998).
RN [14]
RP SEQUENCE FROM N.A.
RX MEDLINE=99383793; PubMed=10456793;
RA Muller K., Tidona C.A., Darai G.;
RT "Identification of a gene cluster within the genome of Chilo
RT iridescent virus encoding enzymes involved in viral DNA replication
RT and processing.";
RL Virus Genes 18:243-264(1999).
RN [15]
RP SEQUENCE FROM N.A.
RX MEDLINE=21342589; PubMed=11448171;
RA Jakob N.J., Muller K., Bahr U., Darai G.;
RT "Analysis of the First Complete DNA Sequence of an Invertebrate
RT Iridovirus: Coding Strategy of the Genome of Chilo Iridescent Virus.";
RL Virology 286:182-196(2001).
RN [16]
RP SEQUENCE FROM N.A.
RA Jakob N.J., Mueller K., Bahr U., Darai G.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF030741; AAK82133.1; -
SQ SEQUENCE 71 AA; 8082 MW; F878E2F1DBE25C6 CRC64;

Query Match 63.5%; Score 33; DB 12; Length 71;
Best Local Similarity 62.5%; Pred. No. 28;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KQAFVFEF 9
Db 11 KQAFIKY 18
|||||
11 KQAFIKY 18

RESULT 12
O93YPI PRELIMINARY; PRT; 420 AA.
AC O93YPI.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Beta-galactosidase.
GN MBK5.28.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucrodis II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,

```

```

RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY059892; AAL24374.1; -
DR EMBL; AY093352; AAM13351.1; -
DR InterPro; IPR001944; Glyco_hydro_35.
DR Pfam; PF01301; Glyco_hydro_35; 1.
DR PRINTS; PR00742; GLHYDRASE35.
SQ SEQUENCE 420 AA; 46206 MW; F75AA65DE16924A6 CRC64;

Query Match 63.5%; Score 33; DB 10; Length 420;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MKQAFVFE 8
Db 67 MQQAYVFE 74
|||||
67 MQQAYVFE 74

RESULT 13
O8XE99 PRELIMINARY; PRT; 431 AA.
AC O8XE99;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Positive and negative sensor protein for pho regulon (Positive and
DE negative sensor protein for pho regulon PhOR).
GN PHOR OR 20498 OR ECS0450.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Rose D.J., Mayhew G.F., Boutin A., Shao Y., Miller L.,
RA Postfai G., Hackett J., Klink S., Lim A., Dimalanta E.T., Potamoudis K.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuwara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [3]
DR EMBL; AE005219; AAG34746.1; -
DR EMBL; AP002551; BAB33873.1; -
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003661; His_kinase.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF02518; HATPase_C; 1.
DR Pfam; PF00512; HSKA; 1.
DR Pfam; PF00989; PAS; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00388; HSKA; 1.
DR SMART; SM00091; PAS; 1.
DR TIGRfams; TIGR00229; sensory_box; 1.
DR PROSITE; PS50109; HIS_KIN; 1.

```

DR PROSITE; PS50112; PAS; 1.
KW Complete proteome.
SQ SEQUENCE 431 AA; 49677 MW; 85C3BFA36D72E25 CRC64;

Query Match 63.5%; Score 33; DB 16; Length 431;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 KQAFVFEFD 10
II I I I I
DB 298 KQTFTEID 306

RESULT 14

O8FKDO PRELIMINARY; PRT; 431 AA.
AC Q8FKDO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Phosphate regulon sensor protein phoR (EC 2.7.3.-).
GN PHOR OR C0509.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AE016756; AAN78987.1; -;
KW Transferase; Complete proteome.
SQ SEQUENCE 431 AA; 49630 MW; 3EE092456D432FFE CRC64;

Query Match 63.5%; Score 33; DB 16; Length 431;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 KQAFVFEFD 10
II I I I I
DB 298 KQTFTEID 306

RESULT 15

O9RHV5 PRELIMINARY; PRT; 436 AA.
AC Q9RHV5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Phosphoglycerate dehydrogenase.
GN PGDA.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MT8148;
RA Kawabata S., Terao Y., Hamada S.;
RT "Molecular cloning, sequence and characterization of a novel
streptococcal phosphoglycerate dehydrogenase gene.";
RL Oral Microbiol. Immunol. 15:58-62(2001).
DR EMBL; AB016077; BAA88823.1; -;
DR InterPro; IPR005289; GTP-binding_dom.
DR InterPro; IPR006073; GTP1_OBG.

DR InterPro; IPR002917; MMR_HSR1.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF01926; MMR_HSR1; 1.
DR PRINTS; PR00326; GTP1OBG.
DR TIGRFAMS; TIGR00650; MG442; 2.
DR TIGRFAMS; TIGR00231; small_GTP; 2.
SQ SEQUENCE 436 AA; 48601 MW; EC52CC8650DED090 CRC64;

Query Match 63.5%; Score 33; DB 2; Length 436;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKOAFVFE 8
: : I I I I I I
DB 417 IROAFVFE 424

Search completed: July 24, 2003, 20:08:38
Job time : 11.5974 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: July 24, 2003, 19:46:15 ; Search time 111.273 Seconds
(without alignments)
205.411 Million cell updates/sec

Title: US-09-546-136-5
Perfect score: 723
Sequence: 1 MELKQAFVEFDENLSSSSG.....LNKMLSTLLGNEENLLQIS 144

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*

2: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*

3: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*

4: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*

5: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*

6: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*

7: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*

8: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*

9: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*

10: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*

11: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*

12: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*

13: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*

14: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*

15: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*

16: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*

17: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*

18: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*

19: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*

20: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*

21: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*

22: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*

23: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

24: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	98	13.6	274	23	ABP73884
2	96.5	13.3	1224	21	ABP18258
3	95	13.1	236	22	ABP61576
4	90.5	12.5	297	19	AAW59151
5	90.5	12.5	453	22	AAW59151
6	90.5	12.5	572	23	ABB09495
7	90.5	12.5	1847	23	ABP73701
8	89.5	12.4	1847	21	AAW5277
9	88.5	12.2	1247	21	ABP18215
					Plasmodium falciparum
					Candida albicans
					Plasmodium falciparum
					Murine prolactin-1
					S. gordonii rggl p
					ERA binding domain
					AmEPV poly(A) poly
					Candida albicans e
					C. albicans CaJ103
					Plasmodium falciparum

10	87	12.0	471	21	AAW54040	Amino acid sequenc
11	86.5	12.0	2539	21	ABP18198	Plasmodium falcipa
12	86	11.9	1948	23	ABP73774	Candida albicans e
13	85.5	11.8	469	19	AAW48742	Human granulocytic
14	85.5	11.8	481	22	ABB60473	Drosophila melanog
15	85	11.8	707	21	AAW50065	Candida albicans p
16	85	11.8	2013	21	ABP18265	Plasmodium falcipa
17	85	11.8	2295	21	ABP18180	Plasmodium falcipa
18	84	11.6	6815	22	ABB66811	Drosophila melanog
19	83	11.5	331	23	ABP38781	Staphylococcus epi
20	82.5	11.4	290	22	ABG16933	Novel human diagno
21	82.5	11.4	367	21	ABW05956	Protein deduced fr
22	82	11.3	314	23	ABB84755	DNA polymerase III
23	82	11.3	1477	20	AAW19981	B. burgdorferi ant
24	82	11.3	1494	20	AAW19980	B. burgdorferi ant
25	81.5	11.3	351	21	ABP18312	Plasmodium falcipa
26	81.5	11.3	1032	23	AAE16107	Human DNAX Toll li
27	81.5	11.3	1040	22	ABB23027	Protein #5026 enco
28	81.5	11.3	1040	22	AAW31141	Peptide #5178 enco
29	81.5	11.3	1040	23	ABG40604	Human peptide enco
30	81.5	11.3	1041	20	AAW41768	Human PRO286 prote
31	81.5	11.3	1041	20	AAW05867	Human Toll protein
32	81.5	11.3	1041	21	AAW44324	Human PRO286 prote
33	81.5	11.3	1041	23	AAO21584	Murine Toll-like r
34	81.5	11.3	1041	23	AAO21586	Murine Toll-like r
35	81.5	11.3	1041	24	ABU61154	Human PRO286 poly
36	81.5	11.3	1059	23	AAO21587	Murine Toll-like r
37	81.5	11.3	1155	23	ABG93108	S. cerevisiae BAX-
38	81.5	11.3	1450	23	ABB55397	Lactococcus lactis
39	81	11.2	446	20	AAW20059	B. burgdorferi ant
40	81	11.2	469	20	AAW20058	B. burgdorferi ant
41	81	11.2	522	22	AAW81755	S. epidermidis ope
42	81	11.2	575	20	AAW83389	Caenorhabditis ele
43	81	11.2	636	22	AAW92772	Human protein sequ
44	80.5	11.1	229	21	AAW41703	Human OREF ORF1467
45	80.5	11.1	331	23	ABB84727	DNA polymerase III

ALIGNMENTS

RESULT 1

ABP73884

ID ABP73884 standard; Protein: 274 AA.

XX

AC ABP73884;

XX

DT 30-JAN-2003 (first entry)

XX

DE Candida albicans essential protein SEQ ID NO 7721.

XX

KW Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;

KW signal transduction; DNA replication; cell division; growth;

KW proliferation; Candida albicans; fungicide; antifungal.

XX

OS Candida albicans.

XX

PN WO200253728-A2.

XX

PD 11-JUL-2002.

XX

PF 26-DEC-2001; 2001WO-US49486.

XX

PR 29-DEC-2000; 2000US-259128P.

PR 20-FEB-2001; 2001US-0792024.

PR 22-AUG-2001; 2001US-314050P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;

XX

DR WPI: 2002-566694/60.

DR N-PSDB: AB232434.

XX Constructing strains for identifying gene products as effective targets
 PT for therapeutic intervention, by inactivating in the strain one allele
 PT of a gene and placing other allele of the gene under conditional
 PT expression -
 XX
 XX Claim 44; SEQ ID NO 7721; 167pp + Sequence Listing; English.
 PS
 XX
 XX The invention relates to constructing (M1) a strain of diploid fungal
 CC cells in which both alleles of a gene are modified, comprising modifying
 CC one allele by insertion or replacement by a cassette having an
 CC expressible selectable marker and modifying other allele by
 CC recombination, of a promoter replacement fragment with a heterologous
 CC promoter, so that expression of the second allele is regulated by the
 CC promoter. (M1) is useful for constructing a strain of diploid fungal
 CC cells in which both alleles of a gene are modified. The diploid fungal
 CC cells having both alleles modified are useful for identifying a gene that
 CC is essential to the survival or growth of a fungus, a gene that
 CC contributes to the virulence and/or pathogenicity of a fungus, a gene
 CC that contributes to the resistance of a diploid fungus to an antifungal
 CC agent, an antifungal agent that inhibits the growth of a diploid fungus
 CC and for identifying a therapeutic agent for treatment of a diploid fungus
 CC disease. (M1) is useful for identifying a compound which modulates the
 CC activity of a gene product, preferably enzymatic activity, carbon
 CC compound catabolism, biosynthetic, transporter, transcriptional,
 CC translational, signal transduction, DNA replication and cell division
 CC activity. The method is useful for identifying a compound having the
 CC ability to inhibit growth or proliferation of *C. albicans* cells and for
 CC treating infection by *C. albicans*. The present sequence is that of an
 CC essential *Candida albicans* protein used in the method of the invention.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification but is based on sequence information supplied to Derwent by
 CC the European Patent Office.
 XX
 XX

Sequence 274 AA;

Query Match 13.6%; Score 98; DB 23; Length 274;
 Best Local Similarity 29.6%; Pred. No. 0.073;
 Matches 32; Conservative 19; Mismatches 41; Indels 16; Gaps 4;
 QY 25 EKVKQNSPNYD-FRITFDIGLYLNKSGVILDYDLKNVLSLVALKRDYLSLSNN 83
 Db 80 EKPKEIFSNKSYEFELNRRG---SNKYTKIENEFDSKNII-----SKDLTIN 126
 QY 84 KOIKKFNKHNKLNKFNLYVINEDIEKPRITKNGILEEVLNKNMLLS 131
 Db 127 SFLKKFEMVNWYKLNKFOFEITNTTSKEGVNNGYFEQV--NKNVIT 172

RESULT 2

AAB18258
 ID AAB18258 standard; Protein: 1224 AA.

XX
 AC AAB18258;

XX 07-NOV-2000 (first entry)

XX Plasmodium falciparum chromosome 2 related protein SEQ ID NO:115.

XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
 KW antimalarial; malaria; protozoacide; infection; insecticide.

XX Plasmodium falciparum.

XX WO200025728-A2.

XX 11-MAY-2000.

XX 05-NOV-1999; 99WO-US26796.

XX 05-NOV-1998; 98US-0107131.

XX (HOFF/) HOFFMAN S.

PA (CARU/) CARUCCI D.
 PA (GARD/) GARDNER M.
 PA (VENT/) VENTER J C.

PI Hoffman S, Carucci D, Gardner M, Venter JC;

XX WPI: 2000-365347/31.

XX
 PT Proteins encoded by chromosome 2 of the human malarial parasite,
 PT Plasmodium falciparum, useful as antimalarial vaccines and in the
 PT diagnosis of *P.falciparum* infection -
 XX
 XX

PS Disclosure; Page 277-280; 577pp; English.

XX The present invention describes proteins and their fragments (I) encoded
 CC by chromosome 2 of the human malarial parasite, plasmodium falciparum.
 CC Also described are: (I) nucleotide sequences (II) encoding (I); and (2)
 CC vaccines against *P. falciparum* infection comprising (I) or (II).
 CC (1) and (II) are useful for the development of vaccines against
 CC *P. falciparum* infection. (I) and polyclonal antisera or a monoclonal
 CC antibody raised to immunogens comprising the sequences of (I), are
 CC useful in the detection of infection with *P. falciparum*. Furthermore,
 CC (I) (especially when they are rifins or secreted or membrane proteins)
 CC can aid the identification of drugs to treat or prevent *P. falciparum*
 CC infection, or they can be used to identify drug resistance in
 CC *P. falciparum*. Sequencing of the Plasmodium chromosome 2 and the
 CC subsequent identification of proteins encoded by it will help to expand
 CC our understanding of parasite biology, a process hampered by the
 CC complexity of the parasitic lifecycle, and provide new targets for
 CC vaccine and drug development. Parasite resistance to drugs and mosquito
 CC resistance to insecticides have led to a resurgence of malaria in many
 CC parts of the world, and there is a pressing need for vaccines and new
 CC drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide
 CC and protein sequences given in the present invention, but which are not
 CC specifically mentioned within the specification.
 XX

Sequence 1224 AA;

Query Match 13.3%; Score 96.5; DB 21; Length 1224;
 Best Local Similarity 31.1%; Pred. No. 0.79;
 Matches 46; Conservative 15; Mismatches 60; Indels 27; Gaps 8;

QY 11 FDENLSSSSGSIHLEKVKQNCSPNYDYFKI-TFIDGILYIKN----KSGVILDK---YDL 62

Db 50 FPNVYONKSHINSEYVN-----NMDYLSLHTGIEKYKRYKNNNNVKNMILKDEDILDY 104

QY 63 KNVVISLVALKRD---YLS---LSSLNNKOIKFKFNKN-----KHLKNKFNLYVINEDI 110

Db 105 NTHLSNHLINHDINFIYSSNNIFNLCNNKNNPKYFPNKSNSNEIKKHOHKNVNTNHIY 164

QY 111 EKRRITKNGILE--EVLNKNMLLSILLGN 136

Db 165 HTKKKNKFNYSNPTEVNVNLSLLSNLKHN 192

RESULT 3

AAB61576
 ID AAB61576 standard; Protein: 236 AA.

XX
 AC AAB61576;

XX 03-APR-2001 (first entry)

XX Murine prolactin-like protein E.

XX Murine prolactin-like protein E; haematopoietic progenitor cell;
 KW placental glycoprotein hormone; proliferation; differentiation; platelet;
 KW megakaryocyte.

XX Mus sp.

XX WO200100788-A2.

PD 04-JAN-2001.
 XX
 PF 23-JUN-2000; 2000WO-US17345.
 XX
 PR 25-JUN-1999; 99US-0141060.
 PR 29-OCT-1999; 99US-0162472.
 XX
 PA (NOUN) UNIV NORTHWESTERN.
 XX
 XX Lefebvre P, Lin J, Cohen I, Linzer D;
 PI WPI; 2001-137950/14.
 XX N-PSDB; AAF28654.
 DR
 DR Modulating survival and differentiation of mammalian hematopoietic
 XX progenitor cells in vitro, involves maintaining the cells in nutritive
 PT medium containing murine prolactin-like protein E or F -
 XX
 PS Claim 2; Fig 11; 45pp; English.
 XX
 CC The present invention relates to a method for maintaining mammalian
 CC haematopoietic progenitor cells (HPC) in vitro by maintaining the cells
 CC in a nutritive medium comprising a placental glycoprotein hormone of the
 CC murine prolactin family. The prolactin hormone can be either murine
 CC prolactin-like protein E (mPLP-E; the present sequence) or murine
 CC prolactin-like protein F (mPLP-F). The method is useful for maintaining
 CC and inducing proliferation, differentiation or both of mammalian
 CC haematopoietic progenitor cells in vitro into megakaryocytes and for
 CC inducing mammalian megakaryocytes into platelets.
 XX
 SQ Sequence 236 AA:
 Query Match 13.1%; Score 95; DB 22; Length 236;
 Best Local Similarity 28.3%; Pred. No. 0.12;
 Matches 47; Conservative 21; Mismatches 44; Indels 54; Gaps 9;
 QY 1 MELKQAFVE-----FDENLSSSGSIHLEKVKQKSPNYDYFKITFDIGVLYIKKSG 54
 DB 34 MELRIFTISEMSAKLIDKFLSSSS-----SDSYDQFMLEFL-GQOELTQNL 81
 QY 55 VILDKYDLK-----NVISLVALKRDYLSLSNNKQIKKFNKIKHLKKNFNL-- 103
 DB 82 TYCHKYSIKVPEDIEEAQNVISL-----EDFPILLSR-----MQAWNETLKNRLNSE 130
 QY 104 -----YVNEIDIEKRTKNGILEEVILNKNMILLGNEML 140
 DB 131 GTPGIDDDILPIYKNIETKIAE--LLED---SKSILSQAYGATENV 171
 RESULT 4
 AAW59151
 ID AAW59151 standard; Protein: 297 AA.
 AC AAW59151;
 XX
 XX 13-AUG-1998 (first entry)
 DT
 DE S. gordonii rggl protein.
 XX
 KW Salt-inducible promoter; lactic acid; food industry; food-grade inducer;
 KW fermentation processes; cheese production.
 XX
 OS Streptococcus gordonii.
 XX
 PN WO9810080-A1.
 XX
 PD 12-MAR-1998.
 XX
 PF 20-AUG-1997; 97WO-EP04755.
 XX
 PR 13-MAR-1997; 97EP-0200744.
 PR 05-SEP-1996; 96EP-0202444.
 XX

PA (UNIL) UNILEVER NV.
 PA (UNIL) UNILEVER PLC.
 XX
 PI Kok J, Ledeboer AM, Sanders JW, Venema G;
 XX WPI; 1998-193629/17.
 DR
 XX
 PT Salt-inducible promoter - derived from lactic acid bacteria, used
 PT for the production of polypeptides in food
 XX
 PS Example 1.4; Fig 7; 111pp; English.
 XX
 CC This sequence represents the rggl protein from Streptococcus gordonii
 CC which is used in a method which identifies and isolates a salt-inducible
 CC promoter (SIP). Using this SIP, salt can be used as a food-grade inducer
 CC in food fermentation processes, e.g. in the production of cheese,
 CC dressings, water-containing spreads, sausages, or sour dough.
 XX
 SQ Sequence 297 AA:
 Query Match 12.5%; Score 90.5; DB 19; Length 297;
 Best Local Similarity 24.5%; Pred. No. 0.48;
 Matches 39; Conservative 31; Mismatches 54; Indels 35; Gaps 8;
 QY 5 QAFVFEEDENLSSSGSIHLEKVKQKSPNYDYFK--ITFDIGVLYIKKSGVILDKYDL 62
 DB 103 EAMAEFFPEKKNYKLTIVIRATLTSCNPDYQVSKGDIEFLDYLSFVEWG----RYEL 158
 QY 63 ---KNVISLVALK-----RDYLSLSNNKQIKKFN-----IKNRHLK--N 99
 DB 159 WLFTNSVNLTLTLETFASEMINRTQFYNNLPENRRRIKMLLNVSACIENNHQVAM 218
 QY 100 KFNLYVINEDEKRTKNGILEEVIL--NKMLLSTLLGN 136
 DB 219 KFLNVIDN---TKIPETDLYDRVLKVKALYSYKGVN 253
 RESULT 5
 AAG99944
 ID AAG99944 standard; Protein: 453 AA.
 XX
 AC AAG99944;
 XX
 DT 27-SEP-2001 (first entry)
 XX
 DE ERA binding domain polypeptide SEQ ID NO 386.
 XX
 KW ERA binding domain; Escherichia coli; GTPase; antimicrobial;
 KW antibacterial; antibiotic; pathogenesis; infection; vaccine;
 KW peptide therapy.
 XX
 OS Escherichia coli.
 XX
 PN WO200153458-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01786.
 XX
 PR 18-JAN-2000; 2000US-0176870.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Lupas AN, Pearce KH;
 XX
 DR WPI; 2001-476108/51.
 XX
 PT New ERA binding domain polypeptides and polynucleotides encoding them,
 PT useful as research reagents and materials for discovery of treatments
 PT and diagnostics for diseases, or for genetic immunisation -
 XX
 PS Claim 1; Page 53; 279pp; English.

XX The present invention relates to ERA binding domain polypeptides
 CC (AAG99559-AAG9989 and AAM00010-AAM00021). The era gene in Escherichia
 CC coli codes for an essential GTPase protein able to autophosphorylate at
 CC serine and/or threonine residues. The protein has potential antimicrobial
 CC and antibacterial activity and is useful in screening for antagonists,
 CC agonists and for compounds with antibiotic activity. The proteins are
 CC also useful in determining their role in pathogenesis of infection,
 CC dysfunction and disease and could be used as part of a vaccine and/or
 CC peptide therapy.

XX Sequence 453 AA;
 SQ
 Query Match 12.5%; Score 90.5; DB 22; Length 453;
 Best Local Similarity 27.6%; Pred. No. 0.85;
 Matches 43; Conservative 23; Mismatches 59; Indels 31; Gaps 7;

QY 13 ENLSSSSGSIHLEKVKONCSPNDYFKITFDGVLVYK-----NKGSGVILD-----KYD 61
 Db 174 EKISQEHKKKSKSVKIACIGKPNVCKSTLINSLLMKKRMITSNKAGTTLDTVLVPIKYN 233

QY 62 LKNVISLVALKRDYLSLSLNKK--QIKKFKNIKKNKLNKFNLYVINEDIEKRITKNGI 119
 Db 234 YKNYIFI-----DTAGMSKKKSKTNKIEKFKIKTLQTEKSHLTLIIIDAKQISKQDL 288

QY 120 LEE-----VILNKM-LLSIL-LGNEENLLO 142
 Db 289 LLSSFTSKSKPLIIVINKCDLLSLKKNLENLIK 324

RESULT 6
 ABB09495
 ID ABB09495 standard; Protein: 572 AA.
 AC ABB09495;
 XX
 XX 01-JUL-2002 (first entry)
 DT
 XX
 DE AnEPV poly(A) polymerase large subunit (AMV038) amino acid sequence.
 XX
 KW AnEPV; gene therapy; viral vector; chromosome mapping; gene mapping;
 KW genetic deficiency disorder; polymerase.
 XX
 OS Ansacta moorei entomopoxvirus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 530..531
 FT /note= "encoded by ATAAAGGC"
 XX
 PN WO200212526-A2.
 XX
 PD 14-FEB-2002.
 XX
 PF 10-AUG-2001; 2001WO-US25287.
 XX
 PR 10-AUG-2000; 2000US-224479P.
 PR 14-SEP-2000; 2000US-0862254.
 XX
 PA (UYFL) UNIV FLORIDA.
 XX
 XX Moyer RW, Li Y, Bawden AL;
 PI
 XX WPI; 2002-227161/28.
 DR N-PSDB; ABL55642.
 DR
 XX Novel recombinant entomopox virus vector useful for delivering
 PT polynucleotide encoding protein to vertebrate cell, comprises
 PT polynucleotide encoding protein operably linked with heterologous
 PT promoter sequence
 XX
 XX Claim 73; Page 123-124; 326pp; English.
 PS
 XX The invention relates to a recombinant entomopox virus (EPV) vector,
 CC

CC comprising a polynucleotide encoding a protein operably linked with a
 CC heterologous promoter sequence. The invention also concerns methods for
 CC providing gene therapy for genetic deficiency disorders. Vectors of the
 CC invention are useful for delivering a polynucleotide encoding a protein
 CC to a vertebrate cell preferably a mammalian cell, such as a human cell.
 CC The vector is introduced into the vertebrate cell by infection in a viral
 CC particle, or by transfection, transduction, or injection either in vitro
 CC or in vivo. The vector is useful for the delivery and expression of
 CC biologically useful proteins in gene therapy protocols, and for
 CC delivering large DNA segments for engineering of vertebrate cells.
 CC Polynucleotides of the invention have applications in techniques such as
 CC their use as insertion sites for foreign genes of interest, hybridisation
 CC probes, for chromosome and gene mapping, in PCR technologies, and in the
 CC production of sense or antisense nucleic acids. Vectors of the invention
 CC provide for stable integration and expression of heterologous DNA in host
 CC cells, and are adapted for accepting large heterologous polynucleotide
 CC inserts which can be delivered in an infected or transformed cell and
 CC expressed in a stable fraction. The current sequence represents the
 CC ansacta moorei entomopoxvirus (AnEPV) poly(A) polymerase large subunit
 CC (AMV038) amino acid sequence.

XX Sequence 572 AA;
 SQ
 Query Match 12.5%; Score 90.5; DB 23; Length 572;
 Best Local Similarity 24.0%; Pred. No. 1.2;
 Matches 36; Conservative 25; Mismatches 52; Indels 37; Gaps 6;

QY 22 IHLEKVKQKSPNDYFKITFDGVLVYKKNKSGVILD-----KYDLKNVISLVALKRDYL- 76
 Db 227 IHLTIGHDTC-----LFSIPFTGHSLSKYKNKFIIDCFDLSNINVKNSLINNIYFI 281

QY 77 --SLSLNN-----KOIKKFKNIKKNKHL--KNKFNLYVINEDIEKRI 114
 Db 282 DPGLOMLNNFRLSENFYSKYIYKMEESLNKYKTLNLYFVNNNNKFNKORLNWLKSDV 341

QY 115 TKNGLILEEVILANKMLLSILLNGNEENLLOIS 144
 Db 342 CRNFPYTTVDNTLILISL-----KELIDIS 366

RESULT 7
 ABB73701
 ID ABB73701 standard; Protein: 1847 AA.
 AC ABB73701;
 XX
 XX 30-JAN-2003 (first entry)
 DT
 XX
 DE Candida albicans essential protein SEQ ID NO 7538.
 XX
 KW Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;
 KW signal transduction; DNA replication; cell division; growth;
 KW proliferation; Candida albicans; fungicide; antifungal.
 XX
 OS Candida albicans.
 XX
 PN WO200253728-A2.
 XX
 PD 11-JUL-2002.
 XX
 PF 26-DEC-2001; 2001WO-US49486.
 XX
 PR 29-DEC-2000; 2000US-259128P.
 PR 20-FEB-2001; 2001US-0792024.
 PR 22-AUG-2001; 2001US-314050P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 XX Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
 PI WPI; 2002-566694/60.
 DR N-PSDB; AB232251.
 DR
 XX

PT Constructing strains for identifying gene products as effective targets
PT for therapeutic intervention, by inactivating in the strain one allele
PT of a gene and placing other allele of the gene under conditional
PT expression -
XX
PS Claim 44: SEQ ID NO 7538; 167pp + Sequence Listing; English.
XX
XX The invention relates to constructing (M1) a strain of diploid fungal
CC cells in which both alleles of a gene are modified, comprising modifying
CC one allele by insertion or replacement by a cassette having an
CC expressible selectable marker and modifying other allele by
CC recombination, of a promoter replacement fragment with a heterologous
CC promoter, so that expression of the second allele is regulated by the
CC promoter. (M1) is useful for constructing a strain of diploid fungal
CC cells in which both alleles of a gene are modified. The diploid fungal
CC cells having both alleles modified are useful for identifying a gene that
CC is essential to the survival or growth of a fungus, a gene that
CC contributes to the virulence and/or pathogenicity of a fungus, a gene
CC that contributes to the resistance of a diploid fungus to an antifungal
CC agent, an antifungal agent that inhibits the growth of a diploid fungus
CC and for identifying a therapeutic agent for treatment of a mammalian
CC disease. (M1) is useful for identifying a compound which modulates the
CC activity of a gene product, preferably enzymatic activity, carbon
CC compound catabolism, biosynthetic, transporter, transcriptional,
CC translational, signal transduction, DNA replication and cell division
CC activity. The method is useful for identifying a compound having the
CC ability to inhibit growth or proliferation of C. albicans cells and for
CC treating infection by C. albicans. The present sequence is that of an
CC essential Candida albicans protein used in the method of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.
XX
SQ Sequence 1847 AA;

Query Match 12.5%; Score 90.5; DB 23; Length 1847;
Best Local Similarity 23.3%; Pred. No. 5.7;
Matches 40; Conservative 25; Mismatches 54; Indels 53; Gaps 7;
QY 19 SGSIHLEKVKQ-----NCSPNYDYFKI-----TFID-----GYL 47
Db 1249 SSSITLNLREYEELEFKLVNCSQPKDLNTVWCNQDFDGLQIDASNVFVDNQASTQAFF 1308
QY 48 YIKNKGVLIDKDYDLKNVLSLVALKRDYLSLSLNKKQ-----IKKFKNIKHKLN- 99
Db 1309 SFINQRLILQYLSLEPHSVKSRTRKREYYSKVLNDKEFVNRTPKVLTFLNLNYSKPNF 1368
QY 100 -----KFNLVYNEDIEKRTKNGILEEVLNKMLLSILLGNEENLL 141
Db 1369 EVQKYEWDQKFNMSLLAEVNAQ--KNGTLDFSVLTKVFR--LLCQTSNLI 1416

RESULT 8
AAY85277
ID AAY85277 standard; Protein; 1847 AA.
XX
AC AAY85277;
XX
DT 04-JUL-2000 (first entry)
XX
DE C. albicans CaJL039 amino acid sequence.
XX
XX CaJL039; essential gene; inhibitor; dihydropteroate synthase; DHPS;
KW 7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase; HPPK; treatment;
KW antimycete substance; infection; fungal infection; mycosis.
XX
OS Candida albicans.
XX
XX WO200015838-A2.
PN
XX 23-MAR-2000.
PD
XX 13-SEP-1999; 99WO-EP07376.

XX
PR 11-SEP-1998; 98EP-0402255.
XX
PA (HMRI) HOECHST MARION ROUSSEL.
XX
PI Lalanne J, Rocher C, Chaltatzis N, Leeuw T, Margerie D, Nitsche A;
PI Reinhard-Rupp J;
XX
DR WPI: 2000-271469/23.
DR N-PSDB: AAA10735.
XX
PT Novel genes from Candida albicans, useful for screening for antifungal
PT drugs for treatment or prevention of fungal infection -
XX
PS Disclosure; Page 93-100; 139pp; English.
XX
CC This sequence represents the protein encoded by the Candida albicans
CC CaJL039 gene. The invention relates to eight nucleotide sequences
CC representing essential C. albicans genes, and the polynucleotides that
CC they encode. An antibody against the polypeptides, and a method for
CC screening for antimycotic substances using one of the essential genes or
CC the peptides encoded by them as a target is also included in the
CC invention. The method may be used to screen for substances that inhibit
CC the expression of the target essential Candida albicans or Aspergillus
CC fumigatus genes or inhibit the activity of essential proteins from these
CC species. This method is particularly useful for screening for inhibitors
CC of 7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase (HPPK) and/or
CC dihydropteroate synthase (DHPS). The method may also be used to identify
CC antimycete substances acting on other mycete species. These inhibitors
CC may be used to treat infections caused by these organisms or may be used
CC prophylactically to prevent infection in immunosuppressed individuals.
CC The nucleotide sequences may be used to identify functionally similar
CC genes in other mycete species, and may also be used as a probe to screen
CC a C. albicans genomic library. The genes and polypeptides may also be
CC used to identify fungal infections. The new method of screening for
CC antimycotic substances using essential genes or proteins allows the
CC identification of drugs which will only inhibit growth of specific mycete
CC species.
XX
SQ Sequence 1847 AA;

Query Match 12.4%; Score 89.5; DB 21; Length 1847;
Best Local Similarity 23.3%; Pred. No. 7.2;
Matches 40; Conservative 25; Mismatches 54; Indels 53; Gaps 7;
QY 19 SGSIHLEKVKQ-----NCSPNYDYFKI-----TFID-----GYL 47
Db 1249 SSSITLNLREYEELEFKLVNCSQPKDLNTVWCNQDFDGLQIDASNVFVDNQASTQAFF 1308
QY 48 YIKNKGVLIDKDYDLKNVLSLVALKRDYLSLSLNKKQ-----IKKFKNIKHKLN- 99
Db 1309 SFINQRLILQYLSLEPHSVKSRTRKREYYSKVLNDKEFVNRTPKVLTFLNLNYSKPNF 1368
QY 100 -----KFNLVYNEDIEKRTKNGILEEVLNKMLLSILLGNEENLL 141
Db 1369 EVQKYEWDQKFNMSLLAEVNAQ--KNGTLDFSVLTKVFR--LLCQTSNLI 1416

RESULT 9
AAB18215
ID AAB18215 standard; Protein; 1247 AA.
XX
AC AAB18215;
XX
DT 07-NOV-2000 (first entry)
XX
DE Plasmodium falciparum chromosome 2 related protein SEQ ID NO:72.
XX
KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KW antimalarial; malaria; protozoacide; infection; insecticide.
XX
OS Plasmodium falciparum.

```

XX XX
KW KW SceI endonuclease; phosphodiester bond; yeast; genetic engineering;
KW KW PCR; polymerase chain reaction; gene cloning.
XX XX
OS OS Saccharomyces cerevisiae.
XX XX
XX XX EP972836-A2.
XX XX
XX XX 19-JAN-2000.
XX XX
XX XX 21-MAY-1999; 99EP-0110008.
XX XX
XX XX 22-MAY-1998; 98JP-0141861.
XX XX
XX XX (RIKA ) INST PHYSICAL & CHEM RES.
XX XX
XX XX Morishima N, Shibata T, Mizumura H;
XX XX
XX XX WPI; 2000-099856/09.
XX XX
XX XX N-PSDB; AA237082.
XX XX
XX XX New modified endonuclease capable of recognizing specific nucleotide
XX XX sequence, useful for genetic engineering techniques -
XX XX
XX XX Disclosure: Fig 1; 43pp; English.
XX XX
XX XX The present sequence represents a yeast SceI endonuclease. The
XX XX endonuclease hydrolyses the phosphodiester bond of a polynucleotide
XX XX chain. The endonuclease is the 50 kDa subunit of a yeast endonuclease.
XX XX The endonuclease gene was modified so that it could be
XX XX mass-produced in an expression system such as E. coli or yeast. The
XX XX endonuclease polynucleotide sequence was modified by substituting
XX XX codons that are unique to mitochondria (the gene is expressed in
XX XX mitochondria) with universal codons. The endonuclease cuts molecules
XX XX within sequence AA237078, and is therefore useful in genetic engineering
XX XX techniques such as PCR (polymerase chain reaction) for cloning,
XX XX amplifying and analysing genes.
XX XX
XX XX Sequence 471 AA;
XX XX
XX XX Query Match 12.0%; Score 87; DB 21; Length 471;
XX XX Best Local Similarity 28.6%; Pred. NO. 2.1;
XX XX Matches 30; Conservative 18; Mismatches 33; Indels 24; Gaps 4;
XX XX
XX QY 32 SPNYDYFKITFDIGLYIKNKGIVLDKYD-----LKNVISIVALKRDYLSL 78
XX Db 86 SRNIIDYKLLYL--FYIILNKIKIEDNTNNNNNISLKYNELLNINLNKLSNIEL 143
XX
XX QY 79 SLSNKQIKFKFNKHNKHLKFNLYVINEDIKRTKNGILEEV 123
XX Db 144 NLSNNFY-----LIDKYLINKYIKYL---DILNIIPNNYIFNNI 179
XX
XX RESULT 11
XX ID AAB18198
XX ID AAB18198 standard; Protein; 2539 AA.
XX XX
XX XX AAB18198;
XX XX
XX DT 07-NOV-2000 (first entry)
XX XX
XX XX Plasmodium falciparum chromosome 2 related protein SEQ ID NO:55.
XX DE
XX DE Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
XX KW antimalarial; malaria; protozoacide; infection; insecticide.
XX KW
XX XX Plasmodium falciparum.
XX OS
XX OS WO200025728-A2.
XX PN
XX PN 11-MAY-2000.
XX PD
XX PD 05-NOV-1999; 99WO-US26796.
XX XX

```

```

XX PR 05-NOV-1998; 98US-0107131.
XX PA (HOFF/) HOFFMAN S.
XX PA (CARU/) CARUCCI D.
XX PA (GARD/) GARDNER M.
XX PA (VENT/) VENTER J C.
XX PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX WPT: 2000-365347/31.
XX PT Proteins encoded by chromosome 2 of the human malarial parasite,
XX PT Plasmodium falciparum, useful as antimalarial vaccines and in the
XX PT diagnosis of P.falciparum infection
XX PS Disclosure: Page 126-133; 577pp; English.
XX CC The present invention describes proteins and their fragments (I) encoded
XX CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
XX CC Also described are: (I) nucleotide sequences (II) encoding (I); and (2)
XX CC vaccines against P. falciparum infection comprising (I) or (II).
XX CC (I) and (II) are useful for the development of vaccines against
XX CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal
XX CC antibody raised to immunogens comprising the sequences of (I), are
XX CC useful in the detection of infection with P. falciparum. Furthermore,
XX CC (I) (especially when they are rifins or secreted or membrane proteins)
XX CC can aid the identification of drugs to treat or prevent P. falciparum
XX CC infection, or they can be used to identify drug resistance in
XX CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
XX CC subsequent identification of proteins encoded by it will help to expand
XX CC our understanding of parasite biology, a process hampered by the
XX CC complexity of the parasitic lifecycle, and provide new targets for
XX CC vaccine and drug development. Parasite resistance to drugs and mosquito
XX CC resistance to insecticides have led to a resurgence of malaria in many
XX CC parts of the world, and there is a pressing need for vaccines and new
XX CC drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide
XX CC and protein sequences given in the present invention, but which are not
XX CC specifically mentioned within the specification.
XX SQ Sequence 2539 AA;
Query Match 12.0%; Score 86.5; DB 21; Length 2539;
Best Local Similarity 26.7%; Pred. No. 23;
Matches 36; Conservative 26; Mismatches 36; Indels 37; Gaps 8;
Qy 36 DYFKITFDIGVLYI-----KNSGVILD---KYDLKNVLSLVALKRDYLSLSLNKKQIKK 88
Db 617 NYSSQFISLLISPFESKNTKLCNLNKHYSYKTKNMN-----NDYTNKYIINKQKNIF 671
Qy 89 FKNIKKHLKKNFN-LYVINEDIEKR-----ITKNGILEEVILNKMLL----- 130
Db 672 YNNIRN-NIKYIRLYNISHQERKKKKKKLTFKKYMLKRECLKNSILNKLIIHPDCKK 730
Qy 131 -SILIG-----NEEN 139
Db 731 GTMILNQNIHNEEN 745
RESULT 12
ID ABP73774 standard; Protein: 1948 AA.
XX AC ABP73774;
XX DT 30-JAN-2003 (first entry)
XX DE Candida albicans essential protein SEQ ID NO 7611.
XX KW Fungus: yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
XX KW signal transduction; DNA replication; cell division; growth;
XX KW proliferation; Candida albicans; fungicide; antifungal.
XX

```

```

OS Candida albicans.
XX PR WO200253728-A2.
XX PD 11-JUL-2002.
XX PF 26-DEC-2001; 2001WO-US49486.
XX PR 29-DEC-2000; 2000US-259128P.
XX PR 20-FEB-2001; 2001US-079202A.
XX PR 22-AUG-2001; 2001US-314050P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
XX WPT: 2002-566694/60.
XX DR N-PSDB; AB232324.
XX CC Constructing strains for identifying gene products as effective targets
XX CC for therapeutic intervention, by inactivating in the strain one allele
XX CC of a gene and placing other allele of the gene under conditional
XX CC expression
XX PS Claim 44; SEQ ID NO 7611; 167pp + Sequence Listing; English.
XX CC The invention relates to constructing (M1) a strain of diploid fungal
XX CC cells in which both alleles of a gene are modified, comprising modifying
XX CC one allele by insertion or replacement by a cassette having an
XX CC expressible selectable marker and modifying other allele by
XX CC recombination, of a promoter replacement fragment with a heterologous
XX CC promoter, so that expression of the second allele is regulated by the
XX CC promoter. (M1) is useful for constructing a strain of diploid fungal
XX CC cells in which both alleles of a gene are modified. The diploid fungal
XX CC cells having both alleles modified are useful for identifying a gene that
XX CC is essential to the survival or growth of a fungus, a gene that
XX CC contributes to the virulence and/or pathogenicity of a fungus, a gene
XX CC that contributes to the resistance of a diploid fungus to an antifungal
XX CC agent, an antifungal agent that inhibits the growth of a diploid fungus
XX CC and for identifying a therapeutic agent for treatment of a mammalian
XX CC disease. (M1) is useful for identifying enzymatic activity, carbon
XX CC compound catabolism, biosynthetic, transporter, transcriptional,
XX CC translational, signal transduction, DNA replication and cell division
XX CC activity. The method is useful for identifying a compound having the
XX CC ability to inhibit growth or proliferation of C. albicans cells and for
XX CC treating infection by C. albicans. The present sequence is that of an
XX CC essential Candida albicans protein used in the method of the invention.
XX CC Note: The sequence data for this patent is not represented in the printed
XX CC specification but is based on sequence information supplied to Derwent by
XX CC the European Patent Office.
XX SQ Sequence 1948 AA;
Query Match 11.9%; Score 86; DB 23; Length 1948;
Best Local Similarity 22.3%; Pred. No. 18;
Matches 43; Conservative 27; Mismatches 67; Indels 56; Gaps 5;
Qy 2 ELKQA-----EVFFDENLSSSGSIHLEKVKQKNSPNYD----- 37
Db 1750 ELKQAHGELDFLKKHLENGREDSEAIKTELQSKMSTSFDIRDOQKLNELLVTKKEFNS 1809
Qy 38 -----FKITFDIGVLYI-----IKNSGVILDKYDLKNVLSLVALKR 73
Db 1810 LVKTNKELNKLKVSDELEKLYSNEQLKYWESKVDLTLSKALDGLAKNEKHEADTKIKLQRSI 1869
Qy 74 DYLSLSLNKKOIKKKFNKKNKHLNKNFNLY-----VINEDIKRTKNGTLEEVILNK 127
Db 1870 KOLEIRVENESQSKRYNDENFDYQNKYHNHKTSTIDIIHNENIEKDLQKLSIQRENI--E 1927
Qy 128 MLLSILLGNEENL 140
Db 1928 MKESMLMLQKEVL 1940

```

RESULT 13

AAW48742
ID AAW48742 standard; Protein; 469 AA.

XX AC AAW48742;

XX DT 01-OCT-1998 (first entry)

XX XX Human granulocytic ehrlichiosis isolate NCH-1 E7 protein.

XX DE Human; granulocytic ehrlichiosis; HGE; aohGE; diagnosis: vaccine;
KW detection; infection.

XX XX Homo sapiens.

XX PN WO9814584-A2.

XX PD 09-APR-1998.

XX PF 30-SEP-1997; 97MO-US17675.

XX PR 01-OCT-1996; 96US-0027180.

XX PA (UYA) UNIV YALE.

XX PI Barthold SW, Fikrig E, Ijdo J, Sun W;

XX DR WPI; 1998-260999/23.

XX DR N-PSDB; AAV18455.

XX PT New isolated human granulocyte ehrlichiosis DNA and proteins - used
PT to develop products for detection, treatment and prevention of human
PT granulocytic ehrlichiosis and related disorders

XX PS Claim 12; Fig 2; 159pp; English.

XX CC The present sequence represents a human granulocytic ehrlichiosis (HGE)
CC isolate NCH-1 E7 protein. The present invention describes HGE DNA
CC molecules and proteins, and methods which can be used for the detection,
CC treatment and prevention of HGE and related disorders caused by
CC infection by protein and DNA molecules which encode them, from the
CC agent of HGE (aohGE). An antibody which binds to an HGE protein can
CC be used in vaccines.

XX SQ Sequence 469 AA;

Query Match 11.8%; Score 85.5; DB 19; Length 469;

Best Local Similarity 27.9%; Pred. No. 2.9;

Matches 43; Conservative 21; Mismatches 51; Indels 39; Gaps 9;

OY 2 ELKQAFVFFDENL-SSSGSIHLEKVKONCSPNYDFKI-----TFIDGYLY-IKN 51

DB 44 ELVRLNFNYTEGIRSKNNKLYIFYKNNCOYLVEQKIDSPKSNVETLI--YFYEIKE 101

OY 52 KSGVLDKDYDLKN-VISLVALKRDYLSLSLNKK-----QIKKFKNIK----- 93

DB 102 ----TYDNOELKNFLYLALENLHISIKQNLGSKLTTELEIPKFNLSKAEQEPIN 157

OY 94 -NKLKNKFNLYVINEIDIEKRITKNGILEEVILN 126

DB 158 QNKRLKD----YQINEKSLREFLNKHQDEIIN 187

RESULT 14

ABB60473
ID ABB60473 standard; Protein; 481 AA.

XX AC ABB60473;

XX DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 8211.

XX XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL04576.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions

XX PS Disclosure; SEQ ID NO 8211; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

XX CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 481 AA;

Query Match 11.8%; Score 85.5; DB 22; Length 481;

Best Local Similarity 24.6%; Pred. No. 3;

Matches 33; Conservative 29; Mismatches 45; Indels 27; Gaps 6;

OY 8 VFEFDENLSSSGSIHLEKVKONCSPNYD---YFKITFTDGYLYIKNKGVL---DKYD 61

DB 8 VSEDELOISNYGIGDCDYEHQDSSDKLDSDFYQILLINLFLYLSNVQGEVVSAPESYD 67

OY 62 L-----KNVISLVALKRDYLSLSLNKKQIKKFNKKNKHLKNFNLYVINEIDIEKRITKN 117

DB 68 FAISESQSLKLRKETHV-----NN--LQSYKKVLRKHL-----QKIRRLKKD 110

OY 118 GILEEVILNKLMLLS 131

DB 111 LGLEQIAVSQMLLT 124

RESULT 15

AA95065
ID AA95065 standard; Protein; 707 AA.

XX AC AA95065;

XX DT 23-JUN-2000 (first entry)

XX DE Candida albicans polypeptide sequence # 33.

KW Candida albicans infection; growth; survival; medicament; AIDS;
KW vulvovaginitis; immunocompromised patient; treat.
XX

Search completed: July 24, 2003, 20:05:59
Job time : 113.273 secs

Result No.	Query			Length	DB	ID	Description
	Score	Match					
1	95	13.1	236	3	US-09-602-848-2		Sequence 12, Appli
2	90.5	12.5	297	3	US-09-068-195-12		Sequence 12, Appli
3	83.5	11.5	210	4	US-09-107-532A-4250		Sequence 4250, Ap
4	83	11.1	331	4	US-09-134-001C-3626		Sequence 3626, Ap
5	80.5	11.1	540	1	US-08-286-325A-2		Sequence 2, Appli
6	79.5	11.0	183	4	US-09-134-001C-3411		Sequence 3411, Ap
7	79	10.9	990	2	US-08-645-1938-15		Sequence 15, Appl
8	79	10.9	1588	5	PCT-US93-07261-11		Sequence 11, Appl
9	79	10.9	1663	5	PCT-US93-07261-16		Sequence 16, Appl
10	77	10.7	411	4	US-09-328-352-6424		Sequence 6424, Ap
11	77	10.7	1101	4	US-09-770-170-8		Sequence 8, Appli
12	76.5	10.6	903	1	US-08-021-601-12		Sequence 12, Appl
13	76.5	10.6	903	1	US-08-082-8498-12		Sequence 12, Appl
14	76.5	10.6	903	5	PCT-US94-01624-12		Sequence 12, Appl
15	76.5	10.6	965	4	US-09-437-277-3		Sequence 3, Appli
16	76	10.5	180	4	US-09-134-001C-4215		Sequence 4215, Ap
17	75.5	10.4	1964	2	US-08-790-912-3		Sequence 3, Appli
18	75.5	10.4	2052	2	US-08-790-912-2		Sequence 2, Appli
19	75	10.4	476	3	US-09-316-083-3		Sequence 3, Appli
20	75	10.4	476	4	US-09-933-700-3		Sequence 3, Appli
21	75	10.4	794	4	US-09-417-485D-8		Sequence 8, Appli
22	74.5	10.3	319	4	US-09-134-001C-5553		Sequence 5553, Ap
23	74.5	10.3	329	4	US-09-071-035-368		Sequence 368, App
24	74.5	10.3	357	4	US-09-071-035-366		Sequence 366, App
25	74.5	10.3	425	4	US-09-134-001C-3654		Sequence 364, Ap
26	74.5	10.3	793	4	US-09-107-532A-6223		Sequence 6223, Ap
27	74.5	10.3	3978	4	US-09-914-259-11		Sequence 11, Appl

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4250:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 210 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...210
; SEQUENCE DESCRIPTION: SEQ ID NO: 4250:
US-09-107-532A-4250

Query Match 11.5%; Score 83.5; DB 4; Length 210;
Best Local Similarity 23.8%; Pred. No. 0.22;
Matches 34; Conservative 28; Mismatches 51; Indels 31; Gaps 5;

Qy 21 SIHLEKVKQKSPNYDYFKITFDGVLGIYIKNKGSGVILDKYDLKNVISLV----- 69
Db 15 SINKKKIKQWN-----FSISEKD-FIIYTKSGS--SGKTTLLNNLSLIDKQYTGNLFPF 66

Qy 70 -----ALKRDYLSLSLNKKOIKKFKNIKNKHLKNKFNLYVINEDIERITKNGI 119
Db 67 EKKVSKHDIQRLRRNHISYLFQNYALLENOTVOYNFELAKKFNSTNSNDHIYNLLTDFSL 126

Qy 120 LEEVILNKMLLSILLGNEENLLOI 143
Db 127 PKDILKEKIPL--LSGGEQOORIAL 148

RESULT 4
US-09-134-001C-3626
; Sequence 3626, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHY
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3626
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3626

Query Match 11.5%; Score 83; DB 4; Length 331;
Best Local Similarity 20.3%; Pred. No. 0.46;
Matches 36; Conservative 34; Mismatches 39; Indels 68; Gaps 8;

Qy 3 LKQAFVF---EPDENLSSSGSTHLEKVKQKSPNYDYFKITFDGVLGIYIKNKGSGVILDK 59
Db 75 VKNSYVFTGKFSKDLNHSDEL-----IK-----FLEK 103

Qy 60 YDLKNVISLV-----KRDYLSLSNNKOIKKFKNIKNKHLK-----NKF----- 101
Db 104 YDGENLIIEVYOPKLDERKKLTPTKKNAQLAKKIQMSEKELKHKKIKNLTNNYKDIKO 163

Qy 102 -----NLYVINEDIERITKNGILEVILNKMLLSILLGN--EENLQITS 144
Db 164 DALELFELTGYNVYVSEGLKLIIFIG--EPTINKKIDILIIINRSLEONYFLLT 218

```

```

: CURRENT APPLICATION NUMBER: US/09/134,001C
: CURRENT FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 3411
: LENGTH: 183
: TYPE: PRT
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3411

Query Match      11.0%; Score 79.5; DB 4; Length 183;
Best Local Similarity 22.7%; Pred. No. 0.48;
Matches 32; Conservative 27; Mismatches 51; Indels 31; Gaps 5

Qy 2 ELKQAFVEFDENLSSSCSITHLEKVKQNCSPNYDYFKITFDGVLTKNKGIVLDKYD 61
Db 30 DLSKGYSQIESHNASPMETFLNLIEVLGTASDFFREPSDEKVLV-KKREQTYDEYD 88
Qy 62 LKNVIS-LVALKROY-----LSLSNNKKQIKFKNKNKLNKNFNLVINDEIKRITK 116
Db 89 KGYILNMLVANSNEEDMEPLILTRPNASYKNFKPSESDTF-----IYCLNGEV----- 137
Qy 117 NGILEEVLNKMLLSILGNE 137
Db 138 -----SLQLGNQ 144

RESULT 7
US-08-645-193B-15
: Sequence 15, Application US/08645193B
: Patent No. 5962253
: GENERAL INFORMATION:
: APPLICANT: Kupke, Thomas
: APPLICANT: Gotz, Friedrich
: APPLICANT: Kempter, Christoph
: APPLICANT: Jung, Gunther
: TITLE OF INVENTION: Oxidative Decarboxylation of Peptides
: TITLE OF INVENTION: Catalyzed by Flavoprotein EpId
: NUMBER OF SEQUENCES: 70
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
: STREET: 1100 New York Avenue, Suite 600
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/645,193B
: FILING DATE: 13-MAY-1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Esmond, Robert W.
: REGISTRATION NUMBER: 32,893
: REFERENCE/DOCKET NUMBER: 0652.1540000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 990 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: not relevant
: MOLECULE TYPE: protein
US-08-645-193B-15

```

Query Match 10.9%; Score 79; DB 2; Length 990;

Best Local Similarity 24.3%; Pred. No. 5.7;

Matches 37; Conservative 25; Mismatches 46; Indels 44; Gaps 8;

QY 2 ELKQAFVFEFDENLSSSGSIHLEKVKQCNPNYDFKITFDGYLYIKNKSGVILDKYD 61

Db 315 ELKQSLA-----DNISEAAILWL-----LSPNH-----PCTKTIRNYHEFFMDKYG 356

QY 62 LKNVISLVAL-----KRDYLSLSNNKQTKKFKNI-----KNKHLKNKFNLYVIN 107

Db 357 FEQVLNKLKLLDSINGFGYPKKD--SYSPFNIAFLKELYLAIONNSHIE-----IT 407

QY 108 EDIEKRITNGILEEV---ILNKMILLSILGN 136

Db 408 ENDVKNLEKNNTVSKINAPVSTIEIYFGN 439

RESULT 8

PCT-US93-07261-11

; Sequence 11, Application PC/TUS9307261

; GENERAL INFORMATION:

; TITLE OF INVENTION: PFEMP3 MALARIA ANTIGEN, ANALOGS, ANTIBODIES AND USES THEREOF

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: John H. C. Blasdale

; STREET: One Giralda Farms

; CITY: Madison

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07940-1000

; COMPUTER READABLE FORM:

; COMPUTER: Apple Macintosh

; OPERATING SYSTEM: Macintosh 6.0.5

; SOFTWARE: Microsoft Word 5.1a

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/07261

; FILING DATE: 19930805

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/927,531

; FILING DATE: 07-AUG-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Blasdale, John H. C.

; REGISTRATION NUMBER: 31,895

; REFERENCE/DOCKET NUMBER: DX0288K

; TELEPHONE: 201-822-7398

; TELEFAX: 201-822-7039

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1588 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US93-07261-11

Query Match

Best Local Similarity 10.9%; Score 79; DB 5; Length 1588;

Matches 34; Conservative 13; Mismatches 28; Indels 22; Gaps 7;

QY 50 KNKSGVILDKY---DLKNVISLVALKRDYLSLSLN-----NKQIKKFKNIKNKHLKNKFN 102

Db 1165 KGSEGLKENVYTNNDLKN-----NDIQNKDLSNKMKNKELLN-KDISNKMKNK-- 1213

QY 103 LYVINEDIKRTKN-GILEEVILNKMILLSILGN 138

Db 1214 -ELLNKLDSNEDMKKNKELLNKRNDKLSI--GNME 1247

RESULT 9

PCT-US93-07261-16

; Sequence 16, Application PC/TUS9307261

; GENERAL INFORMATION:

; TITLE OF INVENTION: PFEMP3 MALARIA ANTIGEN, ANALOGS, ANTIBODIES AND USES THERE

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: John H. C. Blasdale

; STREET: One Giralda Farms

; CITY: Madison

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07940-1000

; COMPUTER READABLE FORM:

; COMPUTER: Apple Macintosh

; OPERATING SYSTEM: Macintosh 6.0.5

; SOFTWARE: Microsoft Word 5.1a

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/07261

; FILING DATE: 19930805

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/927,531

; FILING DATE: 07-AUG-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Blasdale, John H. C.

; REGISTRATION NUMBER: 31,895

; REFERENCE/DOCKET NUMBER: DX0288K

; TELEPHONE: 201-822-7398

; TELEFAX: 201-822-7039

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1663 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; ORIGINAL SOURCE:

; ORGANISM: Plasmodium falciparum

; STRAIN: Malayan Camp

PCT-US93-07261-16

Query Match

Best Local Similarity 10.9%; Score 79; DB 5; Length 1663;

Matches 34; Conservative 13; Mismatches 28; Indels 22; Gaps 7;

QY 50 KNKSGVILDKY---DLKNVISLVALKRDYLSLSLN-----NKQIKKFKNIKNKHLKNKFN 102

Db 1165 KGSEGLKENVYTNNDLKN-----NDIQNKDLSNKMKNKELLN-KDISNKMKNK-- 1213

QY 103 LYVINEDIKRTKN-GILEEVILNKMILLSILGN 138

Db 1214 -ELLNKLDSNEDMKKNKELLNKRNDKLSI--GNME 1247

RESULT 10

US-09-328-352-6424

; Sequence 6424, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBAC

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 6424

; LENGTH: 411

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-6424

Query Match

Best Local Similarity 10.7%; Score 77; DB 4; Length 411;

Matches 29; Conservative 24; Mismatches 40; Indels 24; Gaps 5;

;; FILING DATE: 25-JUN-1993
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/021,601
;; FILING DATE: 12-FEB-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Weber, Kenneth A.
;; REGISTRATION NUMBER: 31,677
;; REFERENCE/DOCKET NUMBER: 15280-161-1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 903 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-082-849B-12

Query Match 10.6%; Score 76.5; DB 1; Length 903;
Best Local Similarity 20.6%; Pred. No. 9.4;
Matches 42; Conservative 33; Mismatches 50; Indels 79; Gaps 11;
QY 7 FVFEFDENLSS-----SSGSIH--LEKVKONCSPN-----YDYFKIT----- 41
Db 550 FDFNFDOOTSQNIKNQAEINATNITVLDKIKLNKAKMILIRDRFHYDRNNAVGADE 609
QY 42 -----FIDGYLY--IKNKG---VILDKYDLKNVISLV 69
Db 610 SVVKEAHREVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRYDMLNSSL- 668
QY 70 ALKRDYLSLSLNKQIKFKNIK--HLKN---KFNLYVINEDIKRI---TKNGIL 120
Db 669 -----RODGGTFIDFKKYNKPLIYISNPNYKVNVAVT--ENTIINPSENGDT 716
QY 121 EEVILNKMILLSILLGNEENLLOIS 144
Db 717 STNGIKKILKVVGLCKGGTVELT 740

RESULT 14
PCT-US94-01624-12
;; Sequence 12, Application PC/TUS9401624
;; GENERAL INFORMATION:
;; APPLICANT: Lepplia, Stephen H.
;; APPLICANT: Klimpel, Kurt R.
;; APPLICANT: Arora, Naveen
;; APPLICANT: Singh, Yogendra
;; APPLICANT: Nichols, Peter J.
;; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
;; RELATED METHODS
;; NUMBER OF SEQUENCES: 31
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW
;; STREET: Steuart Street Tower, 20th Floor, One Market
;; CITY: San Francisco
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94105
;; COMPUTER READABLE FORM:
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US94/01624
;; FILING DATE: June 25, 1993
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Weber, Kenneth A.
;; REGISTRATION NUMBER: 31,677

;; REFERENCE/DOCKET NUMBER: 15280-115
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 543-9600
;; TELEFAX: (415) 543-5043
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 903 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
PCT-US94-01624-12

Query Match 10.6%; Score 76.5; DB 5; Length 903;
Best Local Similarity 20.6%; Pred. No. 9.4;
Matches 42; Conservative 33; Mismatches 50; Indels 79; Gaps 11;
QY 7 FVFEFDENLSS-----SSGSIH--LEKVKONCSPN-----YDYFKIT----- 41
Db 550 FDFNFDOOTSQNIKNQAEINATNITVLDKIKLNKAKMILIRDRFHYDRNNAVGADE 609
QY 42 -----FIDGYLY--IKNKG---VILDKYDLKNVISLV 69
Db 610 SVVKEAHREVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRYDMLNSSL- 668
QY 70 ALKRDYLSLSLNKQIKFKNIK--HLKN---KFNLYVINEDIKRI---TKNGIL 120
Db 669 -----RODGGTFIDFKKYNKPLIYISNPNYKVNVAVT--ENTIINPSENGDT 716
QY 121 EEVILNKMILLSILLGNEENLLOIS 144
Db 717 STNGIKKILKVVGLCKGGTVELT 740

RESULT 15
US-09-437-277-3
;; Sequence 3, Application US/09437277
;; Patent No. 644447
;; GENERAL INFORMATION:
;; APPLICANT: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
;; TITLE OF INVENTION: POLYMER GRAFTING BY POLYSACCHARIDE SYNTHASES
;; FILE REFERENCE: 5820-551
;; CURRENT APPLICATION NUMBER: US/09/437,277
;; CURRENT FILING DATE: 1999-11-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: WordPerfect 8.0 (saved in ASCII format)
;; SEQ ID NO 3
;; LENGTH: 965
;; TYPE: PRT
;; ORGANISM: Pasteurella multocida
US-09-437-277-3

Query Match 10.6%; Score 76.5; DB 4; Length 965;
Best Local Similarity 25.5%; Pred. No. 10;
Matches 28; Conservative 27; Mismatches 44; Indels 11; Gaps 4;
QY 35 YDYFKITFID-GYLYIKKSGVILDKYDLKNVISLVALKRDYLSL-----SNKQIKKF 89
Db 664 YNDRFDDLSRKYIFNKTAEYQEBMDMLKDLKLIQNKDAKIAVSIFYPNTLNGLVKKL 723
QY 90 KNKKNKLNKFNLYVINEDIKRIKNGILEVT--LNKMLLSILLGNE 137
Db 724 NNI-----IEYNKNIFVILLHVDKNHLPDRIKKEILAFYHKHQNILLNND 769

Search completed: July 24, 2003, 20:09:25
Job time : 40.3377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:08:50 ; Search time 61.7143 seconds
(without alignments)
277.106 Million cell updates/sec

Title: US-09-546-136-5

Perfect score: 723

Sequence: 1 MELKQAFVFEFDENLSSSG.....LNKMLLSILLGNEENLLQIS 144

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82.5	11.4	367	9	US-09-861-451A-46
2	81.5	11.3	1032	11	US-09-950-041-37
3	81.5	11.3	1040	9	US-09-864-761-38325
4	81.5	11.3	1041	10	US-09-978-295A-498
5	81.5	11.3	1041	10	US-09-978-697-498
6	81.5	11.3	1041	10	US-09-978-192A-498
7	81.5	11.3	1041	10	US-09-999-832A-498
8	81.5	11.3	1041	11	US-09-978-189-498
9	81.5	11.3	1041	11	US-09-978-608A-498
10	81.5	11.3	1041	11	US-09-978-585A-498
11	81.5	11.3	1041	11	US-09-978-191A-498
12	81.5	11.3	1041	11	US-09-978-403A-498
13	81.5	11.3	1041	11	US-09-978-564A-498
14	81.5	11.3	1041	11	US-09-999-833A-498
15	81.5	11.3	1041	11	US-09-981-915A-498

16	81.5	11.3	1041	11	US-09-978-824-498	Sequence 498, App
17	81.5	11.3	1041	11	US-09-918-585A-498	Sequence 498, App
18	81.5	11.3	1041	11	US-09-978-423A-498	Sequence 498, App
19	81.5	11.3	1041	11	US-09-978-193A-498	Sequence 498, App
20	81.5	11.3	1041	11	US-09-999-830A-498	Sequence 498, App
21	81.5	11.3	1041	11	US-09-978-757A-498	Sequence 498, App
22	81.5	11.3	1041	11	US-09-978-187B-498	Sequence 498, App
23	81.5	11.3	1041	11	US-09-954-987B-184	Sequence 184, App
24	81.5	11.3	1041	11	US-09-954-987B-186	Sequence 186, App
25	81.5	11.3	1041	11	US-09-978-843A-498	Sequence 498, App
26	81.5	11.3	1041	12	US-09-978-375A-498	Sequence 498, App
27	81.5	11.3	1041	15	US-10-095-627-3	Sequence 3, Appli
28	81.5	11.3	1041	15	US-10-017-081A-498	Sequence 498, App
29	81.5	11.3	1041	15	US-10-167-749-498	Sequence 498, App
30	81.5	11.3	1041	15	US-10-013-921A-498	Sequence 498, App
31	81.5	11.3	1041	15	US-10-013-929A-498	Sequence 498, App
32	81.5	11.3	1041	15	US-10-016-177A-498	Sequence 498, App
33	81.5	11.3	1041	15	US-10-235-767-3	Sequence 3, Appli
34	81.5	11.3	1041	15	US-10-166-709A-498	Sequence 498, App
35	81.5	11.3	1059	11	US-09-954-987B-187	Sequence 187, App
36	81.5	11.3	1155	10	US-09-801-368-304	Sequence 304, App
37	81	11.2	575	9	US-09-220-091-7	Sequence 7, Appli
38	81	11.2	1805	11	US-09-820-843A-73	Sequence 73, Appli
39	80.5	11.1	322	15	US-10-012-140-20	Sequence 20, Appli
40	80	11.1	2167	10	US-09-801-368-56	Sequence 56, Appli
41	79.5	11.0	979	10	US-09-801-368-284	Sequence 284, App
42	77.5	10.7	1174	15	US-10-205-841-40	Sequence 40, Appli
43	77	10.7	1101	15	US-10-059-962-8	Sequence 8, Appli
44	76.5	10.6	965	11	US-09-842-484A-2	Sequence 2, Appli
45	76.5	10.6	965	11	US-09-842-484A-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-861-451A-46
; Sequence 46, Application US/09861451A
; Patent No. US20020068289A1
; GENERAL INFORMATION:
; APPLICANT: Commonwealth Scientific & Industrial Research Orga
; TITLE OF INVENTION: Methods of Identifying Antigen Gene Sequences
; FILE REFERENCE: FF34033/01
; CURRENT APPLICATION NUMBER: US/09/861,451A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PP7273
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Deduced protein
; OTHER INFORMATION: sequence from clone PAD922
US-09-861-451A-46

Query Match	11.4%	Score 82.5;	DB 9;	Length 367;
Best Local Similarity	25.5%;	Pred. No. 2.1;		
Matches	36;	Conservative	27;	Mismatches 67;
				Indels 11;
				Gaps 5;
QY	4	KQAFVFEFDENLSSSGSIHLEKVKQNCSPNDYFKITFDIGLYIKNKGVLQYDKL 63		
DB	109	KQAKVEFDYRSKDYEDOLKIKOEKA---FFIKTLNVKALMLKKEAQLEIDKFTAQ 164		
QY	64	NVISLVAKRDYLSLSLNNKQIKKFKNLKHLKNLYVINEDIEKRTKNGTLEEV 123		
DB	165	NNLTSYIDERNYKFIANNKALVT-TDLKNYSFFKK--QALINKEIAFLDORNL--- 218		
QY	124	ILNKMLLSL-LGNEENLQI 143		
DB	219	LLEKNLFSFLNISETEKLFEI 239		

RESULT 2
US-09-950-041-37
Sequence 37, Application US/09950041
Publication No. US20030032090A1
GENERAL INFORMATION:
APPLICANT: Hardman, Gerard T.
APPLICANT: Rock, Fernando L.
APPLICANT: Bazan, J. Fernando
APPLICANT: Kastelein, Robert A.
APPLICANT: Ho, Stephen W.K.
APPLICANT: Liu, Yong-Jun
TITLE OF INVENTION: RECEPTOR PROTEINS; RELATED REAGENTS AND METHODS
FILE REFERENCE: DX0724XX1
CURRENT APPLICATION NUMBER: US/09/950,041
CURRENT FILING DATE: 2002-05-06
PRIORITY APPLICATION NUMBER: 09/728,540
PRIORITY FILING DATE: 2000-11-28
PRIORITY APPLICATION NUMBER: 60/207,558
PRIORITY FILING DATE: 2000-05-25
PRIORITY APPLICATION NUMBER: 09/073,363
PRIORITY FILING DATE: 1999-06-05
PRIORITY APPLICATION NUMBER: 60/044,293
PRIORITY FILING DATE: 1997-05-07
PRIORITY APPLICATION NUMBER: 60/072,212
PRIORITY FILING DATE: 1998-01-22
PRIORITY APPLICATION NUMBER: 60/076,947
PRIORITY FILING DATE: 1998-03-05
NUMBER OF SEQ ID NOS: 45
SOFTWARE: Patentin version 3.1
SEQ ID NO 37
LENGTH: 1032
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: misc_feature
LOCATION: (725)..(725)
OTHER INFORMATION: The 'Xaa' at location 725 stands for Ile.
US-09-950-041-37

Query Match 11.3%; Score 81.5; DB 11; Length 1032;
Best Local Similarity 25.5%; Pred. No. 10;
Matches 38; Conservative 25; Mismatches 61; Indels 25; Gaps 8;
QY 1 MELKQAFVFEFDENLSSSGSIHLEKVKQNCSPNYDYK-----ITFDG-YLYTK 50
Db 59 LQLSDNFTHITN--ESFGQLNLTKINLHNPVQHONGNPGIOSNGLNITDGAFLNLK 116
QY 51 NKSQVILDKYDLKNVISLVALKRDYLSLSNN-----KOIKPKNTKKNKHLKKNFL 103
Db 117 NLRELLLEDNQLPQIPS--GLPESLTELSTLQNNIYNTITKGISRLINKNLKLYL--AWNC 172
QY 104 YVINEDIEKRITKNGILEEVILNKMLLSI 132
Db 173 Y-FNKVCEKTNIEDGVF-ETLTNLELLSL 199

RESULT 3
US-09-864-761-38325
Sequence 38325, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 38325
LENGTH: 1040
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005859.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EST_HUMAN HIT: AW401674.1, EVALUE 2.00e-16
OTHER INFORMATION: EST_HUMAN HIT: AW401674.1, EVALUE 2.00e-16
OTHER INFORMATION: SWISSPROT HIT: 002833, EVALUE 9.00e-22
US-09-864-761-38325

Query Match 11.3%; Score 81.5; DB 9; Length 1040;
Best Local Similarity 25.5%; Pred. No. 10;
Matches 38; Conservative 25; Mismatches 61; Indels 25; Gaps 8;
QY 1 MELKQAFVFEFDENLSSSGSIHLEKVKQNCSPNYDYK-----ITFDG-YLYTK 50
Db 67 LQLSDNFTHITN--ESFGQLNLTKINLHNPVQHONGNPGIOSNGLNITDGAFLNLK 124
QY 51 NKSQVILDKYDLKNVISLVALKRDYLSLSNN-----KOIKPKNTKKNKHLKKNFL 103
Db 125 NLRELLLEDNQLPQIPS--GLPESLTELSTLQNNIYNTITKGISRLINKNLKLYL--AWNC 180
QY 104 YVINEDIEKRITKNGILEEVILNKMLLSI 132
Db 181 Y-FNKVCEKTNIEDGVF-ETLTNLELLSL 207

RESULT 4
US-09-978-295A-498
Sequence 498, Application US/09978295A
Patent No. US20020156006A1
GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurnev, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630PIC11
CURRENT APPLICATION NUMBER: US/09/978,295A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29

PRIOR APPLICATION NUMBER: 60/083554
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083558
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083559
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083500
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083742
 PRIOR FILING DATE: 1998-04-30
 PRIOR APPLICATION NUMBER: 60/084366
 PRIOR FILING DATE: 1998-05-05
 PRIOR APPLICATION NUMBER: 60/084414
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084441
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084637
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084639
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084640
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084598
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084627
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084643
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/085339
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085338
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085323
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085582
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085700
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085689
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085580
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085573
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085704
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085697

Query Match 11.3%; Score 81.5; DB 10; Length 1041;
 Best Local Similarity 25.5%; Pred No. 10;
 Matches 38; Conservative 25; Mismatches 61; Indels 25; Gaps 8;
 QY 1 MELKQAFVEFDENISSSGSHLEKVKONCSPNYDYK-----ITFDIG-YLYTK 50
 DB 68 LLDLSDNFTHITN--ESFQGLQNLTKINLHNPVQHQNGPQIOGSLNITDGAFLNLK 125
 QY 51 NKSQVILDKYDKLVNLSVALKRDYLSLSLNN-----KQKFKTKNKKHKLKNKFN 103
 DB 126 NLRELLEDNOLPQIPS--GLPESLTSLQNNIYNTKIGISRLINKNLYL--AWNC 181
 QY 104 VVINEDIEKRITKNGILEEVLNKMLLSI 132
 DB 182 Y-FNKVCENTNEDGVF-ETLTNLELLSL 208

RESULT 5
 US-09-978-697-498
 : Sequence 498, Application US/09978697
 : Patent No. US20020169284A

GENERAL INFORMATION:
 APPLICANT: Ashkenazi, Avi
 APPLICANT: Baker Kevin P.
 APPLICANT: Botstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Eaton, Dan
 APPLICANT: Ferrara, Napoleon
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Gottlieb, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, J. Christopher
 APPLICANT: Gurney, Austin L.
 APPLICANT: Hillan, Kenneth J.
 APPLICANT: Kljavin, Ivar J.
 APPLICANT: Kuo, Sophia S.
 APPLICANT: Napier, Mary A.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Shelton, David L.
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William I.
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 FILE REFERENCE: P2630PIC27
 CURRENT APPLICATION NUMBER: US/09/978,697
 CURRENT FILING DATE: 2001-10-16
 PRIOR APPLICATION NUMBER: 09/918585
 PRIOR FILING DATE: 2001-07-30
 PRIOR APPLICATION NUMBER: 60/062250
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/064249
 PRIOR FILING DATE: 1997-11-03
 PRIOR APPLICATION NUMBER: 60/065311
 PRIOR FILING DATE: 1997-11-13
 PRIOR APPLICATION NUMBER: 60/066364
 PRIOR FILING DATE: 1997-11-21
 PRIOR APPLICATION NUMBER: 60/077450
 PRIOR FILING DATE: 1998-03-10
 PRIOR APPLICATION NUMBER: 60/077632
 PRIOR FILING DATE: 1998-03-11
 PRIOR APPLICATION NUMBER: 60/077641
 PRIOR FILING DATE: 1998-03-11
 PRIOR APPLICATION NUMBER: 60/077649
 PRIOR FILING DATE: 1998-03-11
 PRIOR APPLICATION NUMBER: 60/077791
 PRIOR FILING DATE: 1998-03-12
 PRIOR APPLICATION NUMBER: 60/078004
 PRIOR FILING DATE: 1998-03-13
 PRIOR APPLICATION NUMBER: 60/078886
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/078936
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/078910
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/078939
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/079294
 PRIOR FILING DATE: 1998-03-25
 PRIOR APPLICATION NUMBER: 60/079656
 PRIOR FILING DATE: 1998-03-26
 PRIOR APPLICATION NUMBER: 60/079664
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/079689
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/079663
 PRIOR FILING DATE: 1998-03-27

;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079920
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/079923
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/080105
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545

;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 11.3%; Score 81.5; DB 10; Length 1041;

Best Local Similarity 25.5%; Pred. No. 10;

Matches 38; Conservative 25; Mismatches 61; Indels 25; Gaps 8;

QY 1 MELKQAFVFEFDENLSSSGSIHLEKVKQKNCSPNYDYFK-----ITFDIG-YLYIK 50
DB 68 LDLSDNFTHITN--ESFQGLNLTKNLNHPNVOHQNGNPGIQSNGLNITDGAFLNLK 125
QY 51 NKSGVLDKYDLKNVLSLVALKROYLSLSLN-----KOIKKFKNKHLKNKPNL 103
DB 126 NURELLDNQLPQIPS--GLPESLTELSIONNYYNITKGISRLINKNLVL--AWNC 181
QY 104 YVINEDIKRTKNGILEEVILNKMLLSI 132
DB 182 Y-FNKVCEKTNIEDGVF-ETLTNLELISL 208

RESULT 6

US-09-978-192A-498

; Sequence 498, Application US/09978192A

Patent No. US2002017753A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630PIC9
CURRENT APPLICATION NUMBER: US/09/978,192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29

```
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
```

```
Query Match 11.3% Score 81.5; DB 10; Length 1041;
Best Local Similarity 25.3%; Pred. No. 10;
Matches 38; Conservative 25; Mismatches 61; Indels 25; Gaps 8;

Qy 1 MELKQAFVFEDEMLSSSGSIHLEKVKQNCSPNYDFK-----ITFDG-YLYIK 50
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 LDLSDNFTHITN--ESFOGLQNLTKLNHNPNVQHONGNPGIQSNGLNITDGAFLNK 125
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 51 NKSGVILDYDLKNVISLVALKRDYLSLSLNN-----KQIKKFKNIKKNHKNENL 103
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 126 NLRELLLEDNQLPQPS--GUPESITELSLQNNIYNTITKGISRLNKLNYL--AWN 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 104 YVINEDEIKRITTKGILEEVLNKLMLLSI 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 Y-FNKVCEKTNIEDGVF-ETLTNLELLSL 208
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 7

US-09-999-832A-498

```
; Sequence 498, Application US/09999832A
; Publication No. US20020192706A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC63
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
```

1	:	PRIOR APPLICATION NUMBER: 60/079666
2	:	PRIOR FILING DATE: 1998-03-27
3	:	PRIOR APPLICATION NUMBER: 60/079728
4	:	PRIOR FILING DATE: 1998-03-27
5	:	PRIOR APPLICATION NUMBER: 60/079786
6	:	PRIOR FILING DATE: 1998-03-27
7	:	PRIOR APPLICATION NUMBER: 60/079920
8	:	PRIOR FILING DATE: 1998-03-30
9	:	PRIOR APPLICATION NUMBER: 60/079923
10	:	PRIOR FILING DATE: 1998-03-30
11	:	PRIOR APPLICATION NUMBER: 60/080105
12	:	PRIOR FILING DATE: 1998-03-31
13	:	PRIOR APPLICATION NUMBER: 60/080107
14	:	PRIOR FILING DATE: 1998-03-31
15	:	PRIOR APPLICATION NUMBER: 60/080165
16	:	PRIOR FILING DATE: 1998-03-31
17	:	PRIOR APPLICATION NUMBER: 60/080194
18	:	PRIOR FILING DATE: 1998-03-31
19	:	PRIOR APPLICATION NUMBER: 60/080327
20	:	PRIOR FILING DATE: 1998-04-01
21	:	PRIOR APPLICATION NUMBER: 60/080328
22	:	PRIOR FILING DATE: 1998-04-01
23	:	PRIOR APPLICATION NUMBER: 60/080333
24	:	PRIOR FILING DATE: 1998-04-01
25	:	PRIOR APPLICATION NUMBER: 60/080334
26	:	PRIOR FILING DATE: 1998-04-01
27	:	PRIOR APPLICATION NUMBER: 60/081070
28	:	PRIOR FILING DATE: 1998-04-08
29	:	PRIOR APPLICATION NUMBER: 60/081049
30	:	PRIOR FILING DATE: 1998-04-08
31	:	PRIOR APPLICATION NUMBER: 60/081071
32	:	PRIOR FILING DATE: 1998-04-08
33	:	PRIOR APPLICATION NUMBER: 60/081195
34	:	PRIOR FILING DATE: 1998-04-08
35	:	PRIOR APPLICATION NUMBER: 60/081203
36	:	PRIOR FILING DATE: 1998-04-09
37	:	PRIOR APPLICATION NUMBER: 60/081229
38	:	PRIOR FILING DATE: 1998-04-09
39	:	PRIOR APPLICATION NUMBER: 60/081955
40	:	PRIOR FILING DATE: 1998-04-15
41	:	PRIOR APPLICATION NUMBER: 60/081817
42	:	PRIOR FILING DATE: 1998-04-15
43	:	PRIOR APPLICATION NUMBER: 60/081819
44	:	PRIOR FILING DATE: 1998-04-15
45	:	PRIOR APPLICATION NUMBER: 60/081952
46	:	PRIOR FILING DATE: 1998-04-15
47	:	PRIOR APPLICATION NUMBER: 60/081838
48	:	PRIOR FILING DATE: 1998-04-15
49	:	PRIOR APPLICATION NUMBER: 60/082568
50	:	PRIOR FILING DATE: 1998-04-21
51	:	PRIOR APPLICATION NUMBER: 60/082569
52	:	PRIOR FILING DATE: 1998-04-21
53	:	PRIOR APPLICATION NUMBER: 60/082704
54	:	PRIOR FILING DATE: 1998-04-22
55	:	PRIOR APPLICATION NUMBER: 60/082804
56	:	PRIOR FILING DATE: 1998-04-22
57	:	PRIOR APPLICATION NUMBER: 60/082700
58	:	PRIOR FILING DATE: 1998-04-22
59	:	PRIOR APPLICATION NUMBER: 60/082797
60	:	PRIOR FILING DATE: 1998-04-22
61	:	PRIOR APPLICATION NUMBER: 60/082796
62	:	PRIOR FILING DATE: 1998-04-23
63	:	PRIOR APPLICATION NUMBER: 60/083336
64	:	PRIOR FILING DATE: 1998-04-27
65	:	PRIOR APPLICATION NUMBER: 60/083322
66	:	PRIOR FILING DATE: 1998-04-28
67	:	PRIOR APPLICATION NUMBER: 60/083392
68	:	PRIOR FILING DATE: 1998-04-29
69	:	PRIOR APPLICATION NUMBER: 60/083495
70	:	PRIOR FILING DATE: 1998-04-29
71	:	PRIOR APPLICATION NUMBER: 60/083496
72	:	PRIOR FILING DATE: 1998-04-29
73	:	PRIOR APPLICATION NUMBER: 60/083499

?	PRIOR FILING DATE:	1998-04-29
?	PRIOR APPLICATION NUMBER:	60/083545
?	PRIOR FILING DATE:	1998-04-29
?	PRIOR APPLICATION NUMBER:	60/083554
?	PRIOR FILING DATE:	1998-04-29
?	PRIOR APPLICATION NUMBER:	60/083558
?	PRIOR FILING DATE:	1998-04-29
?	PRIOR APPLICATION NUMBER:	60/083559
?	PRIOR FILING DATE:	1998-04-29
?	PRIOR APPLICATION NUMBER:	60/083500
?	PRIOR FILING DATE:	1998-04-29
?	PRIOR APPLICATION NUMBER:	60/083742
?	PRIOR FILING DATE:	1998-04-30
?	PRIOR APPLICATION NUMBER:	60/084366
?	PRIOR FILING DATE:	1998-05-05
?	PRIOR APPLICATION NUMBER:	60/084414
?	PRIOR FILING DATE:	1998-05-06
?	PRIOR APPLICATION NUMBER:	60/084441
?	PRIOR FILING DATE:	1998-05-06
?	PRIOR APPLICATION NUMBER:	60/084637
?	PRIOR FILING DATE:	1998-05-07
?	PRIOR APPLICATION NUMBER:	60/084639
?	PRIOR FILING DATE:	1998-05-07
?	PRIOR APPLICATION NUMBER:	60/084640
?	PRIOR FILING DATE:	1998-05-07
?	PRIOR APPLICATION NUMBER:	60/084598
?	PRIOR FILING DATE:	1998-05-07
?	PRIOR APPLICATION NUMBER:	60/084600
?	PRIOR FILING DATE:	1998-05-07
?	PRIOR APPLICATION NUMBER:	60/084627
?	PRIOR FILING DATE:	1998-05-07
?	PRIOR APPLICATION NUMBER:	60/084643
?	PRIOR FILING DATE:	1998-05-07
?	PRIOR APPLICATION NUMBER:	60/085339
?	PRIOR FILING DATE:	1998-05-13
?	PRIOR APPLICATION NUMBER:	60/085338
?	PRIOR FILING DATE:	1998-05-13
?	PRIOR APPLICATION NUMBER:	60/085323
?	PRIOR FILING DATE:	1998-05-13
?	PRIOR APPLICATION NUMBER:	60/085582
?	PRIOR FILING DATE:	1998-05-15
?	PRIOR APPLICATION NUMBER:	60/085700
?	PRIOR FILING DATE:	1998-05-15
?	PRIOR APPLICATION NUMBER:	60/085689
?	PRIOR FILING DATE:	1998-05-15
?	PRIOR APPLICATION NUMBER:	60/085579
?	PRIOR FILING DATE:	1998-05-15
?	PRIOR APPLICATION NUMBER:	60/085580
?	PRIOR FILING DATE:	1998-05-15
?	PRIOR APPLICATION NUMBER:	60/085573
?	PRIOR FILING DATE:	1998-05-15
?	PRIOR APPLICATION NUMBER:	60/085704
?	PRIOR FILING DATE:	1998-05-15
?	PRIOR APPLICATION NUMBER:	60/085697

Query Match	11.3%	Score 81.5;	DB 10;	Length 1041;
Best Local Similarity	25.5%;	Pred. No. 10;		
Matches 38:	Conservative 25;	Mismatches 61;	Indels 25;	Gaps 8;

Qy	1	MELKQAFVFEFDENLSSSSGSIHLEKVKQKQCSN	YDK-----ITFIDG-YLYIK	50
Db	68	LDLSDNFIHTN-ESFOGLNLTKLNNHPNVQHONG	PGIQSGNGLNITDGAFLNLK	125
Qy	51	NKSGVILDKYDLKNVLSLVALKRDYLSLSLNN	-----KQIKKFKNIKNKHUKNFNL	103
Db	126	NLRLELLEDNQLPOIPS-GLPESLTSLIQNNIY	NTREGISRLINLKNLYL-AWNC	181
Qy	104	YVINEDIEKRITKNGILEEVILNKMLLSI	132	
Db	182	Y-FNKYCEKTNIEDGVF-ETLTNLELNL	208	

US-09-978-189-498
: Sequence 498, Application US/09978189
: Publication No. US20030004102A1
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James;
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2630PIC7
: CURRENT APPLICATION NUMBER: US/09/978,189
: CURRENT FILING DATE: 2001-10-15
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: PRIOR APPLICATION NUMBER: 60/078004
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: 60/078886
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078936
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078939
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/079294
: PRIOR FILING DATE: 1998-03-25
: PRIOR APPLICATION NUMBER: 60/079656
: PRIOR FILING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/079664
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079663
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079728
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079786
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079920
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/079923
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/080105
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080107
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080165
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080194
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080327
: PRIOR FILING DATE: 1998-04-01
: PRIOR APPLICATION NUMBER: 60/080328
: PRIOR FILING DATE: 1998-04-01
: PRIOR APPLICATION NUMBER: 60/080333
: PRIOR FILING DATE: 1998-04-01
: PRIOR APPLICATION NUMBER: 60/080334
: PRIOR FILING DATE: 1998-04-01
: PRIOR APPLICATION NUMBER: 60/081070
: PRIOR FILING DATE: 1998-04-08
: PRIOR APPLICATION NUMBER: 60/081049
: PRIOR FILING DATE: 1998-04-08
: PRIOR APPLICATION NUMBER: 60/081071
: PRIOR FILING DATE: 1998-04-08
: PRIOR APPLICATION NUMBER: 60/081195
: PRIOR FILING DATE: 1998-04-08
: PRIOR APPLICATION NUMBER: 60/081203
: PRIOR FILING DATE: 1998-04-09
: PRIOR APPLICATION NUMBER: 60/081229
: PRIOR FILING DATE: 1998-04-09
: PRIOR APPLICATION NUMBER: 60/081955
: PRIOR FILING DATE: 1998-04-15
: PRIOR APPLICATION NUMBER: 60/081817
: PRIOR FILING DATE: 1998-04-15
: PRIOR APPLICATION NUMBER: 60/081819
: PRIOR FILING DATE: 1998-04-15
: PRIOR APPLICATION NUMBER: 60/081952
: PRIOR FILING DATE: 1998-04-15
: PRIOR APPLICATION NUMBER: 60/081838
: PRIOR FILING DATE: 1998-04-15
: PRIOR APPLICATION NUMBER: 60/082568
: PRIOR FILING DATE: 1998-04-21
: PRIOR APPLICATION NUMBER: 60/082569
: PRIOR FILING DATE: 1998-04-21
: PRIOR APPLICATION NUMBER: 60/082704
: PRIOR FILING DATE: 1998-04-22
: PRIOR APPLICATION NUMBER: 60/082804
: PRIOR FILING DATE: 1998-04-22
: PRIOR APPLICATION NUMBER: 60/082700
: PRIOR FILING DATE: 1998-04-22
: PRIOR APPLICATION NUMBER: 60/082797
: PRIOR FILING DATE: 1998-04-22
: PRIOR APPLICATION NUMBER: 60/082796
: PRIOR FILING DATE: 1998-04-23
: PRIOR APPLICATION NUMBER: 60/083336
: PRIOR FILING DATE: 1998-04-27
: PRIOR APPLICATION NUMBER: 60/083322
: PRIOR FILING DATE: 1998-04-28
: PRIOR APPLICATION NUMBER: 60/083392
: PRIOR FILING DATE: 1998-04-29
: PRIOR APPLICATION NUMBER: 60/083495
: PRIOR FILING DATE: 1998-04-29
: PRIOR APPLICATION NUMBER: 60/083496
: PRIOR FILING DATE: 1998-04-29


```

Query Match      11.3%; Score 81.5; DB 11; Length 1041;
Best Local Similarity 25.5%; Pred. No. 10;
Matches 38; Conservative 25; Mismatches 61; Indels 25; Gaps 8;

QY      1 MELKQAFVEFDENLSSSGSIHLEKVKQNCSPNYDFK-----ITFDIG-YLYIK 50
      .  : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      68 LDLSNFIYHITN--ESFQGLQNUTKLNHNPNVQHONGNPGIOGNGINITDGAFLNLK 125
      .  : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      51 NKSQGVLDKYDLKNVILSVALKRDYLSLSNN-----KQIKKFKNKHLKKNKFNL 103
      .  : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      126 NLRELLEDNQLPOIPS--GLPESLTLSLQNNIYINITKEGISRLINLKNLYL--AANC 181
      .  : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      104 YVINEDIEKRTKNGILEEVILNKMMLLSI 132
      .  : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      182 Y-FNKVCEKTNIEDGVF-ETITNLELISL 208
      .  : : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

```

Query Match      11.3%; Score 81.5; DB 11; Length 1041;
Best Local Similarity 25.5%; Pred. No. 10;
Matches 38; Conservative 25; Mismatches 61; Indels 25; Gaps 8;

Qy 1 MELKQAFVFEEFNLSSSSGSIHLEKKVKQNCSPNYDFK-----ITFDG-YLYIK 50
Db 68 LDLSDNPFITHTN--ESFOGLQNLTKLNHNPNVQHONGPGIQSNGLNITDGAFLNLK 125
Qy 51 NKSGVILDKYDLKNVTSLVALKRDYLSLSLN-----KQIKKKFNKHLKKNKFN 103
Db 126 NLRELLLEDNQLPOIPS--GLPSLELSLIQNNIYNITREGISRLINLKNLYL--AWNC 181
Qy 104 YVINEDIKRTKNGILEEVILNKMULSI 132
Db 182 Y-FNKVCEKTNIEDGVF-ETLTNLELSL 208

```

RESULT 10
US-09-978-585A-498
; Sequence 498, Application US/09978585A
; Publication No. US20030049633A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon

```

: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630PIC15
: CURRENT FILING DATE: 2001-10-16
: NUMBER OF SEQ ID NOS: 624
: Prior Application removed - See File Wrapper or Palm
: SEQ ID NO 498
: LENGTH: 1041
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-978-585A-498

Query Match      11.3%; Score 81.5; DB 11; Length 1041;
Best Local Similarity 25.5%; Pred. No. 10;
Matches 38; Conservative 25; Mismatches 61; Indels 25; Gaps 8;

QY 1 MELKQAFVEFDEMLSSSSSIHLEKVKQCSNYDYFK-----ITFIDG-YLYIK 50
DB 68 LLDLSDFNFIHTN--ESFGQLNLTKNLHNHNPNVQHONGNPGIQSNGLNITDGAFLNKL 125
QY 51 NKSCVILDKYLDKNVLSVALKRDYLSLSLSNN-----KQKKKKNKNKHLKKNFNL 103
DB 126 NLRELLEDNQLPOIPS--GLPSLTSLQNNIYNTKIGSRILNKLNLV--AWNC 181
QY 104 YVNEDEKRTKTKNGILEEVILKMLLSI 132
DB 182 Y-FNKVCERTNIEDGVF-ETLTNLELLSL 208

RESULT 11
US-09-978-191A-498
: Sequence 498, Application US/09978191A
: Publication No. US20030050239A1
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker, Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnovers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630PIC15
: CURRENT FILING DATE: 2001-10-16
: NUMBER OF SEQ ID NOS: 624
: Prior Application removed - See File Wrapper or Palm
: SEQ ID NO 498
: LENGTH: 1041
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-978-585A-498

: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630PIC4
: CURRENT FILING DATE: 2001-10-15
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/074450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: PRIOR APPLICATION NUMBER: 60/078004
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: 60/078886
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078936
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078939
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/079294
: PRIOR FILING DATE: 1998-03-25
: PRIOR APPLICATION NUMBER: 60/079656
: PRIOR FILING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/079664
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079663
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079728
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079786
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079920
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/079923
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/080105
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080107
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080165
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080194
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080327
: PRIOR FILING DATE: 1998-04-01
```

;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637

;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697
;;
Query Match 11.3%; Score 81.5; DB 11; Length 1041;
Best Local Similarity 25.5%; Pred. No. 10;
Matches 38; Conservative 25; Mismatches 61; Indels 25; Gaps 8;

OY 1 MELKQAFVEFDENLSSSGSIHLKVKONCSPNYDYFK-----ITFTDG-YLYIK 50
Db 68 LDLSDNFITHITN--ESFQGLQNLTKLNHNPNVYOHQNGNPGIOSNLNITDGAFLNLK 125
OY 51 NKSGVILDKYDLKNVSLVALKRDYLSLSLN-----KOIKFKFNKKNKKNKFN 103
Db 126 NLRELLEDNQLPQIPS--GLPESLTELSONNIYNYITKEGISRLNKLNL--AMNC 181
OY 104 YVINEDIKRTKNGILEEVLNKMLLSI 132
Db 182 Y-FNKVCEKTNIEDGVF-ETLTNLELLSL 208

RESULT 12
US-09-978-403A-498
;; Sequence 498, Application US/09978403A
;; Publication No. US20030050240A1
;; GENERAL INFORMATION:
;; APPLICANT: Ashkenazi, Avi
;; APPLICANT: Baker Kevin P.
;; APPLICANT: Botstein, David
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Eaton, Dan
;; APPLICANT: Ferrara, Napoleon
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Fong, Sherman
;; APPLICANT: Gao, Wei-Qiang
;; APPLICANT: Gerber, Hanspeter
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Hillan, Kenneth J

APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630P1C17
CURRENT FILING DATE: 2002-03-19
CURRENT APPLICATION NUMBER: US/09/978.403A
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06

```

: PRIOR APPLICATION NUMBER: 60/084637
: PRIOR FILING DATE: 1998-05-07
: PRIOR APPLICATION NUMBER: 60/084639
: PRIOR FILING DATE: 1998-05-07
: PRIOR APPLICATION NUMBER: 60/084640
: PRIOR FILING DATE: 1998-05-07
: PRIOR APPLICATION NUMBER: 60/084598
: PRIOR FILING DATE: 1998-05-07
: PRIOR APPLICATION NUMBER: 60/084600
: PRIOR FILING DATE: 1998-05-07
: PRIOR APPLICATION NUMBER: 60/084627
: PRIOR FILING DATE: 1998-05-07
: PRIOR APPLICATION NUMBER: 60/084643
: PRIOR FILING DATE: 1998-05-07
: PRIOR APPLICATION NUMBER: 60/085339
: PRIOR FILING DATE: 1998-05-13
: PRIOR APPLICATION NUMBER: 60/085338
: PRIOR FILING DATE: 1998-05-13
: PRIOR APPLICATION NUMBER: 60/085323
: PRIOR FILING DATE: 1998-05-13
: PRIOR APPLICATION NUMBER: 60/085582
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085700
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085689
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085579
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085580
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085573
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085704
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/085697

```

```

Query Match      11.3%  Score 81.5;  DB 11;  Length 1041;
Best Local Similarity 25.5%;  Pred. No. 10;
Matches 38;  Conservative 25;  Mismatches 61;  Indels 25;  Gaps 8;

Qy 1 MELKQAFVFEDENLSSSSSIHLKVKQKCNPNYDFK-----ITFDG-VLYIK 50
Db 68 LLDLSDNFITHITN--ESFQGLQNLTKLNHNPNVQHONGNPGIQSNGLNITDGAFLNK 125
Qy 51 NKSGVILDKYDLKNVLSVALKRDYLSLSLN-----KQIKKFNKIKNKLKKNFL 103
Db 126 NLRELLLEDNLQFOIPS--GLPESLTSLIQNNIYNITKGISRLINKNLYL--AWNC 181
Qy 104 YVINEDEKTRITKNGILEEVILNKMLLSI 132
Db 182 Y-FNKVCEKTNIEDGVF-ETLTNLELLSL 208

```

RESULT 13

```

US-09-978-564A-498
: Sequence 498, Application US/09978564A
: Publication No. US20030050241A1
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker, Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.

```

```

: APPLICANT: Hillan, Kenneth J
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630PIC25
: CURRENT APPLICATION NUMBER: US/09/978,564A
: CURRENT FILING DATE: 2001-10-16
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: PRIOR APPLICATION NUMBER: 60/078004
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: 60/078886
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078936
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078939
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/079294
: PRIOR FILING DATE: 1998-03-25
: PRIOR APPLICATION NUMBER: 60/079656
: PRIOR FILING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/079664
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079663
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079728
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079786
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079920
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/079923
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/080105
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080107
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080165
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080194
: PRIOR FILING DATE: 1998-03-31

```

;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441

;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 11.3%; Score 81.5; DB 11; Length 1041;
Best Local Similarity 25.5%; Pred. No. 10;
Matches 38; Conservative 25; Mismatches 61; Indels 25; Gaps 8;

QY 1 MELKQAFVFEFENLSSSSGSIHLEKVKQKONCSPNDYDK-----ITFTDG-YLYIK 50
DB 68 LDLSDNFITHITN--ESEFGLQNLTKINLHNPNVQHONGNPGIOGSLNGLITDGAFLNLK 125
QY 51 NKSGVILOKYDLKNVISLVAKRDYLSLSLN-----KOIKFKFNKKNKHLKNK 103
DB 126 NLRELLDNQLPQIPS--GUPESLTELQNNIYNITKGISRLINLKNLYL--AWNC 181
QY 104 YVINEDIKRTKNGILEEVILNKMELSI 132
DB 182 Y-FNKVCEKTNIEDGVF-ETLNLLELSL 208

RESULT 14
US-09-999-833A-498
; Sequence 498, Application US/09999833A
; Publication No. US20030054405A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630PLC65
CURRENT APPLICATION NUMBER: US/09/999,833A
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06

1 PRIOR APPLICATION NUMBER: 60/084441
2 PRIOR FILING DATE: 1998-05-06
3 PRIOR APPLICATION NUMBER: 60/084637
4 PRIOR FILING DATE: 1998-03-07
5 PRIOR APPLICATION NUMBER: 60/084639
6 PRIOR FILING DATE: 1998-05-07
7 PRIOR APPLICATION NUMBER: 60/084640
8 PRIOR FILING DATE: 1998-05-07
9 PRIOR APPLICATION NUMBER: 60/084598
10 PRIOR FILING DATE: 1998-05-07
11 PRIOR APPLICATION NUMBER: 60/084600
12 PRIOR FILING DATE: 1998-05-07
13 PRIOR APPLICATION NUMBER: 60/084627
14 PRIOR FILING DATE: 1998-05-07
15 PRIOR APPLICATION NUMBER: 60/084643
16 PRIOR FILING DATE: 1998-05-07
17 PRIOR APPLICATION NUMBER: 60/085339
18 PRIOR FILING DATE: 1998-05-13
19 PRIOR APPLICATION NUMBER: 60/085338
20 PRIOR FILING DATE: 1998-05-13
21 PRIOR APPLICATION NUMBER: 60/085323
22 PRIOR FILING DATE: 1998-05-13
23 PRIOR APPLICATION NUMBER: 60/085582
24 PRIOR FILING DATE: 1998-05-15
25 PRIOR APPLICATION NUMBER: 60/085700
26 PRIOR FILING DATE: 1998-05-15
27 PRIOR APPLICATION NUMBER: 60/085689
28 PRIOR FILING DATE: 1998-05-15
29 PRIOR APPLICATION NUMBER: 60/085579
30 PRIOR FILING DATE: 1998-05-15
31 PRIOR APPLICATION NUMBER: 60/085580
32 PRIOR FILING DATE: 1998-05-15
33 PRIOR APPLICATION NUMBER: 60/085573
34 PRIOR FILING DATE: 1998-05-15
35 PRIOR APPLICATION NUMBER: 60/085704
36 PRIOR FILING DATE: 1998-05-15
37 PRIOR APPLICATION NUMBER: 60/085697

Query Match 11.3%; Score 81.5; DB 11; Length 1041;

Best Local Similarity 25.5%; Pred. No. 10;

Matches 38; Conservative 25; Mismatches 61; Indels 25; Gaps 8;

QY 1 MELKQAFVEFENJSSSGSHLEKVKQNCSPNDYFK-----ITFDIG-VLYTK 50
DB 68 LBLDNFTHITN--ESFQGLQNLTKVNLNHPNVQHQNGNPGIQSNGNLITDGAFLNKL 125
QY 51 NKSQVTLKDYDLKNTSLVALKRDYLSLSLNN-----KQTKFKNKHLKKNKFL 103
DB 126 NURELLENDNQIPQS--GLPESLTELSQLNNIYNTKGISRLNKLNLVL--AWNC 181
QY 104 YVINEDIKRTKNGILEEVILNKMLLSI 132
DB 182 Y-FNKVCKETNIEDGVF-ETLTNLELLSL 208

RESULT 15

US-09-981-915A-498

Sequence 498, Application US/09981915A

Publication No. US20030054986A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kiljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC12
CURRENT APPLICATION NUMBER: US/09/981,915A
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31


```

? PRIOR APPLICATION NUMBER: 60/080194
? PRIOR FILING DATE: 1998-03-31
? PRIOR APPLICATION NUMBER: 60/080327
? PRIOR FILING DATE: 1998-04-01
? PRIOR APPLICATION NUMBER: 60/080328
? PRIOR FILING DATE: 1998-04-01
? PRIOR APPLICATION NUMBER: 60/080333
? PRIOR FILING DATE: 1998-04-01
? PRIOR APPLICATION NUMBER: 60/080334
? PRIOR FILING DATE: 1998-04-01
? PRIOR APPLICATION NUMBER: 60/081070
? PRIOR FILING DATE: 1998-04-08
? PRIOR APPLICATION NUMBER: 60/081049
? PRIOR FILING DATE: 1998-04-08
? PRIOR APPLICATION NUMBER: 60/081071
? PRIOR FILING DATE: 1998-04-08
? PRIOR APPLICATION NUMBER: 60/081195
? PRIOR FILING DATE: 1998-04-08
? PRIOR APPLICATION NUMBER: 60/081203
? PRIOR FILING DATE: 1998-04-09
? PRIOR APPLICATION NUMBER: 60/081229
? PRIOR FILING DATE: 1998-04-09
? PRIOR APPLICATION NUMBER: 60/081955
? PRIOR FILING DATE: 1998-04-15
? PRIOR APPLICATION NUMBER: 60/081817
? PRIOR FILING DATE: 1998-04-15
? PRIOR APPLICATION NUMBER: 60/081819
? PRIOR FILING DATE: 1998-04-15
? PRIOR APPLICATION NUMBER: 60/081952
? PRIOR FILING DATE: 1998-04-15
? PRIOR APPLICATION NUMBER: 60/081838
? PRIOR FILING DATE: 1998-04-15
? PRIOR APPLICATION NUMBER: 60/082568
? PRIOR FILING DATE: 1998-04-21
? PRIOR APPLICATION NUMBER: 60/082569
? PRIOR FILING DATE: 1998-04-21
? PRIOR APPLICATION NUMBER: 60/082704
? PRIOR FILING DATE: 1998-04-22
? PRIOR APPLICATION NUMBER: 60/082804
? PRIOR FILING DATE: 1998-04-22
? PRIOR APPLICATION NUMBER: 60/082700
? PRIOR FILING DATE: 1998-04-22
? PRIOR APPLICATION NUMBER: 60/082797
? PRIOR FILING DATE: 1998-04-22
? PRIOR APPLICATION NUMBER: 60/082796
? PRIOR FILING DATE: 1998-04-23
? PRIOR APPLICATION NUMBER: 60/083336
? PRIOR FILING DATE: 1998-04-27
? PRIOR APPLICATION NUMBER: 60/083322
? PRIOR FILING DATE: 1998-04-28
? PRIOR APPLICATION NUMBER: 60/083392
? PRIOR FILING DATE: 1998-04-29
? PRIOR APPLICATION NUMBER: 60/083495
? PRIOR FILING DATE: 1998-04-29
? PRIOR APPLICATION NUMBER: 60/083496
? PRIOR FILING DATE: 1998-04-29
? PRIOR APPLICATION NUMBER: 60/083499
? PRIOR FILING DATE: 1998-04-29
? PRIOR APPLICATION NUMBER: 60/083545
? PRIOR FILING DATE: 1998-04-29
? PRIOR APPLICATION NUMBER: 60/083554
? PRIOR FILING DATE: 1998-04-29
? PRIOR APPLICATION NUMBER: 60/083558
? PRIOR FILING DATE: 1998-04-29
? PRIOR APPLICATION NUMBER: 60/083559
? PRIOR FILING DATE: 1998-04-29
? PRIOR APPLICATION NUMBER: 60/083500
? PRIOR FILING DATE: 1998-04-29
? PRIOR APPLICATION NUMBER: 60/083742
? PRIOR FILING DATE: 1998-04-30
? PRIOR APPLICATION NUMBER: 60/084366
? PRIOR FILING DATE: 1998-05-05
? PRIOR APPLICATION NUMBER: 60/084414

```

```

? PRIOR FILING DATE: 1998-05-06
? PRIOR APPLICATION NUMBER: 60/084441
? PRIOR FILING DATE: 1998-05-06
? PRIOR APPLICATION NUMBER: 60/084637
? PRIOR FILING DATE: 1998-05-07
? PRIOR APPLICATION NUMBER: 60/084639
? PRIOR FILING DATE: 1998-05-07
? PRIOR APPLICATION NUMBER: 60/084640
? PRIOR FILING DATE: 1998-05-07
? PRIOR APPLICATION NUMBER: 60/084598
? PRIOR FILING DATE: 1998-05-07
? PRIOR APPLICATION NUMBER: 60/084600
? PRIOR FILING DATE: 1998-05-07
? PRIOR APPLICATION NUMBER: 60/084627
? PRIOR FILING DATE: 1998-05-07
? PRIOR APPLICATION NUMBER: 60/084643
? PRIOR FILING DATE: 1998-05-07
? PRIOR APPLICATION NUMBER: 60/085339
? PRIOR FILING DATE: 1998-05-13
? PRIOR APPLICATION NUMBER: 60/085338
? PRIOR FILING DATE: 1998-05-13
? PRIOR APPLICATION NUMBER: 60/085323
? PRIOR FILING DATE: 1998-05-13
? PRIOR APPLICATION NUMBER: 60/085582
? PRIOR FILING DATE: 1998-05-15
? PRIOR APPLICATION NUMBER: 60/085700
? PRIOR FILING DATE: 1998-05-15
? PRIOR APPLICATION NUMBER: 60/085689
? PRIOR FILING DATE: 1998-05-15
? PRIOR APPLICATION NUMBER: 60/085579
? PRIOR FILING DATE: 1998-05-15
? PRIOR APPLICATION NUMBER: 60/085580
? PRIOR FILING DATE: 1998-05-15
? PRIOR APPLICATION NUMBER: 60/085573
? PRIOR FILING DATE: 1998-05-15
? PRIOR APPLICATION NUMBER: 60/085704
? PRIOR FILING DATE: 1998-05-15
? PRIOR APPLICATION NUMBER: 60/085697

```

```

Query Match      11.3%; Score 81.5; DB 11; Length 1041;
Best Local Similarity 25.5%; Pred. No. 10;
Matches 38; Conservative 25; Mismatches 61; Indels 25; Gaps 8;

QY 1 MELKQAFVEFDENLSSSGSIHLEKVKONCSPNDYFK-----ITFTDG-YLYIK 50
Db 68 LDLSDNFITHITN--ESFQGLQNLTKINLHNPNVQHONGNPGIOSNGLNITDGAFLNLK 125
QY 51 NKSGVILDKYDLKNVISLVALKRDVLSLSLNN-----KOIKKFKNIKHLKNKFNL 103
Db 126 NLRELLLEDNQLPQIPS--GLPESLTELSLIQQNNIYNTKIGISRLINLKNLYL--AWNC 181
QY 104 YVINEDIEKRITKNGILEEVLNKNMLLSI 132
Db 182 Y-FNKVCKEKTNIEDGVF-ETLNLLELSSL 208

Search completed: July 24, 2003, 20:21:04
Job time : 65.7143 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:04:00 ; Search time 46.7532 Seconds
(without alignments)
296.200 Million cell updates/sec

Title: US-09-546-136-5

Perfect score: 723

Sequence: 1 MELKQARVFEDENLSSSG.....LNKMLLSILLGNENLLQIS 144

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109.5	15.1	902	2 D84970	ribonuclease E (im
2	104.5	14.5	1001	2 H64593	type III restricti
3	101	14.0	651	2 F30111	DNA primase (impor
4	98.5	13.6	371	2 A64481	hypothetical prote
5	98.5	13.6	1039	2 S62509	probable vesicular
6	98.5	13.6	1044	2 T50213	probable vesicular
7	96.5	13.3	1224	2 E71611	hypothetical prote
8	95	13.1	404	2 F71714	alanine racemase (
9	94.5	13.1	2510	2 T28160	hypothetical prote
10	94	13.0	2166	2 G70163	hypothetical prote
11	92.5	12.8	329	2 T28412	ORF MSV251 hypothe
12	92	12.7	602	2 E90568	DNA primase (impor
13	92	12.7	2269	2 T18472	hypothetical prote
14	91	12.6	377	2 G69937	lipopolysaccharide
15	91	12.6	1183	2 F90559	conserved hypothet
16	90.5	12.5	214	2 E90512	hypothetical prote
17	90.5	12.5	255	2 F89775	hypothetical prote
18	90.5	12.5	297	2 A41898	positive regulator
19	90.5	12.5	949	2 F90086	chromosomal region
20	89.5	12.4	714	2 C90100	hypothetical prote
21	89.5	12.4	786	2 T18469	hypothetical prote
22	89.5	12.4	3724	2 T18427	hypothetical prote
23	89	12.3	862	2 C97343	hypothetical prote
24	88.5	12.2	1247	2 E71616	hypothetical prote
25	88	12.2	498	2 B97046	hypothetical prote
26	88	12.2	1650	2 T18444	hypothetical prote
27	87.5	12.1	159	2 S06843	olfactory mucosa p
28	87.5	12.1	335	2 S27674	hypothetical prote
29	87.5	12.1	386	2 S58755	ribosomal protein

30 87.5 12.1 407 2 H81289 probable sugar tra
31 87.5 12.1 456 2 A90136 hypothetical prote
32 87.5 12.1 542 2 E90604 hypothetical prote
33 87.5 12.1 693 2 A90110 26S proteasome reg
34 87.5 12.1 789 2 G90587 lipoprotein (impor
35 87.5 12.1 5005 2 F82884 hypothetical prote
36 87 12.0 230 2 H97343 ABC transport syst
37 87 12.0 355 2 H97715 alanine racemase (

ALIGNMENTS

RESULT 1

D84970

ribonuclease E [imported] - Buchnera sp. (strain APS)

C:Species: Buchnera sp.

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001

C:Accession: D84970

R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera s

A:Reference number: A84930; MUID:20445173; PMID:10993077

A:Accession: D84970

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-902 <STO>

A:Cross-references: GB:AP000398; GSPDB:GN00144

A:Experimental source: strain APS

C:Genetics:

A:Gene: rne; BU347

Query Match 15.1%; Score 109.5; DB 2; Length 902;
Best Local Similarity 28.6%; Pred. No. 0.6;
Matches 44; Conservative 26; Mismatches 55; Indels 29; Gaps 6;

Qy 10 EFDEN-LSSSSGSIHLEKVKONCPNVDYFKITFDGVLVIKKN--SGVILDKYDLKNVI 66
Db 544 EAQENILKKNNYNNILLKVLNS-NRNFIKMITWFKNSFFIKNMLITSDFKKNTLKNV- 601
Qy 67 SLVALKRDYLSLSLNKKQIKKF-----KNIKNKHKN-----KFN 102
Db 602 NNIFPKKYSLSLNKKNNQKRVILSKLFEANIENIPLNKKKLDTSANLYLDNIERKKN 661
Qy 103 LYVINEDIEKRTKNGILEEVILNKMILLSILLGN 136
Db 662 ITRKNDLIQKNIHENSYLKVLNMRYNINIINN 695

RESULT 2

H64593

type III restriction enzyme R protein - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999

C:Accession: H64593

R:Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann,

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khatalak, H.G.; Glodek, A.; McF

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Katpik, P.D.; Smith, H.O.; Fraser

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: H64593

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A;Residues: 1-371 <BUL>
A;Cross-references: GB:U67585; GB:L77117; NID:g1592088; PID:AAB99458.1; PID:g1592088
C;Genetics:
A;Map position: FOR1419701-1420816

	Query Match	13.6%	Score 98.5;	DB 2;	Length 371;
	Best Local Similarity	27.6%;	Pred. No. 1.6;		
	Matches 32;	Conservative	28;	Mismatches 41;	Indels 15; Gaps 4;
Qy	13	ENLSSSGSHLEKVKQNCSPNYDYFKITIDGYLYIKKSGVILDKYLDKKNVISLVALK..72			
Db	175	KNLIISAGGINSPIRLKMKMIDDENIGKNLEFDFTVTV..--GGILEDSYLNKDISMLVYKK 231			
Qy	73	RDVLSLSLNKK-----QIKK-FKNKKK-----HLKKNFNLYIVINEDEKRIKT 116			
Db	232	YKNFMPLATYSKLLINEIKDYQKDVKEKDVAGIMIKIKDENNGSVLNDVYKKKEIT 287			

RESULT 5
S62509
S62509 probable vesicular transport factor - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 21-Jan-2000
C:Accession: S62509; T38504
R:Jones, L.; Murphy, L.; McNeill, A.; Simpson, I.; Harris, D.
submitted to the EMBL Data Library, October 1995
A:Reference number: S62507
A:Accession: S62509
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1039 <ION>
A:Cross-references: EMBL:266525; NID:gl044926; PIDN:CAA91425.1; PID:gl044929
R:Jones, L.; Murphy, L.; McNeill, A.; Simpson, I.; Harris, D.; Barrell, B.G.; Rajand
submitted to the EMBL Data Library, October 1995
A:Reference number: 221798
A:Accession: T38504
A:Status: preliminary; translated from GB/EMBL/DD8J
A:Molecule type: DNA
A:Residues: 1-1039 <J02>

```

Query Match          13.6%; Score 98.5; DB 2; Length 1039;
Best Local Similarity 27.9%; Pred. No. 5.1;
Matches 39; Conservative 27; Mismatches 43; Indels 31; Gaps 8;

Qy 22 IHLEKVKQNCSPNDYFKTFIDCYL-----YTKKSGVILDKYDKNVLISLVALK-RDY 75
Db 673 IELDYTKSNCKNQBEEMQV-LREGHSEIKDFTEHSESKLTQQLDDIKNOFGIISSKNRDL 731
Ov 76 LSLSLSNKKNTKKFKNT-----KNKHKLNKENLYVINEPIEDYKRT-----KNGILE 121

```

Qy	122	EVILKNMLLSILIGNENLL	141
		: : :	
Db	782	EAELSKALNDNLGNENII	801

RESULT 6
T50213

probable vesicular transport factor [imported] - fission yeast (Schizosaccharomyces
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
C:Accession: T50213
R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Murphy, L.; Jones, L.
submitted to the EMBL Data Library, January 2000
A:Reference number: Z25046
A:Accession: T50213

DNA primase [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: E90568
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galissou, F.; Moszer, J.
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis
A:Reference number: A99512; MUID:113267165; PMID:11353084
A:Accession: E90568
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-602 <WUR>
A:Cross-references: GB:AL445566; PID:g14089867; PIDN:CA13626.1; GSPDB:GN001153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPU_4530
A:Genetic code: SGC3

Query Match	12.7%	Score 92	DB 2	Length 602	
Best Local Similarity	26.7%	Pred. No. 8.8			
Matches	40	Conservative 24	Mismatches 48	Indels 38	Gaps 8
Qy	12	DENLSSSSGIHLEKVKONCSPNDYFKITFDGILYIKNKGSGVILDKYDLKNVISLVAL	71		
Db	338	DELLKQKGNFLKELIEKNKKPAID-----YLY-----EFLTSKYINNSEHLISF	384		
Qy	72	KRDYLS--LSLSNNKQIKKFKNKNKHLKN--KFNLYV-----INE--DIEKR-----	114		
Db	385	VNDFAPLLSSONNIVIDKYENLLMKHSINIRKFKLTQYONKYDRINQYNSLQORFYEPKQ	444		
Qy	115	-----TKNGILEEVILN---KMLLSILLGN	136		
Db	445	KKANSREELTKKQYQNLVYKLLLSLNS	474		

hypothetical protein C0440c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence-revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T18472
R:Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z18937
A:Accession: T18472
A:Status: preliminary; translated from GP/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-2269 <LAW>
A:CROSS-references: EMBL:AL008970; NID:e1407852; PID:e1332566; PIDN:CAAL5615.1
C:Genetics:
A:Gene: C0440c
A:Map position: 3

```

Query Match      12.7% Score 92; DB 2; Length 2269;
Best Local Similarity 24.9%; Pred. No. 40;
Matches 47; Conservative 27; Mismatches 57; Indels 58; Gaps 9;

Qy   11 FDENLSSSGSIHLEKVKQN-----CSPNYD-----YFKI 40
     | : | | | | : | | | | : | | |
Db   1812 YECNDESSMCSNNKSKOKKKCNLFNRSVLGFLSMCPFDTKRNIYDNQTADIIAVC 1871

Qy   41 TFDGVLVIK-----NKSGVLDKYDLKNVISLV-----ALKROYLSLSLS 81
     | : | | | | : | | | | : | | |
Db   1872 SSID-YSYTNDTTWNVYTVLPNPKFLTFELSSIPYHDNVIIQTITNLTSZDLSLKIF 1930

Qy   82 NNKOI--KFEKNIKLNKLKNFYLVINEDIK-RITKNGILE-----EVLNKMJLSILL 134
     | : | | | | : | | | | : | | |
Db   1931 DSKEIILKNMKMKHKKNKTNTTYDTNSTINPNPNIITHKDLLRVNLYRYVMKLLSLTL 1990

Qy   135 G-NEENLLQ 142
     | : | | | |

```

	Best Local Similarity	30.08; Pred. No. 23;	Matches	Conservative	23; Mismatches	55; Indels	28; Gaps
Qy	12	DENLSSSGSIHLEKVKQSPNDYDKITF-----TDGYLYI---KNKSVIDLK	59	: : :	:	:	:
Db	557	NNNSFKSGNLFLLKLNTVFSNYLSVGYSNLKHLMELDKILVPVEKSIAGNIULNK	616	:	:	:	:
Qy	60	YLKINV-----ISAVALKROYLSLSLSNNKOIKKEFNKTKHLKNKPNLYINVEDIEKKI	114	:: :::::	:	:	:
Db	617	YLKINFNFEEISSLSYEIVQNWKLLNVNLI--FEINKN----NKYNKKIF--DIENKM	669	:	:	:	:
Qy	115	TKGILBEVLNKMMLSLTILGNEENLLOIS	144	:	:	:	:
Db	670	DS-----LSTYVDNKNKLISLL--DEEFUKLN	695	: :	:	:	:

Search completed: July 24, 2003, 20:10:26
Job time : 50.7532 secs

Query Match	12.6%	Score 91	DB 2	Length 377
Best Local Similarity	23.0%	Pred. No. 6.2		
Matches	37	Conservative	31	Mismatches 43
			Indels	50
			Gaps	8

QY	7	FVFEFDENLSSSGSIHLEKVKONCSPNYDYFK-----	ITFIDGYLYIKKNSGVILD	58
		: : : : : : : :		
Db	140	FAIESSDRVTVAVSALAAE-----	TYOLIKPEKKIETIYNFIDERVYLYKNTAAIKE	191
		: : : : : : : :		
QY	59	KYDLKNVYISLVALKRDYLSLSNNKQIKK-----	FKNIKNKHLKNKNLKNVINEDIE	111
		: : : : : : : : :		
Db	192	KHGI-----	LPDEKVVITHVSNFRKVKRVQDVIRVFRNIAG---KTKAKLLLVGDGPE	240
		: : : : : : : : :		
QY	112	KR-----	ITKNGLIEEVILNKMGLSILLNGNE---ENLLQIS	144
		: : : : : : : : :		
Db	241	KSTACELIRKYGLLEDQV-----	LMIGNODRVEDLYSIS	273
		: : : : : : : : :		

RESULT 15
F90559
conserved hypothetical protein MYPU_3820 [imported] - Mycoplasma pulmonis (strain UAB CT1P)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: F90559
R:Chambaud, J.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: F90559
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1183 <KUR>
A:Cross-references: GB:AL445566; PID:g14089796; PIDN:CAC13555.1; GSPDB:GN00153
A:Experimental source: strain UAB CT1P
C:Genetics:
A:Gene: MYPU_3820
A:Genetic code: SGC3

Query Match
12.6%; Score 91; DB 2; Length 1183;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 19:47:29 ; Search time 25.2468 Seconds

(without alignments)
268.226 Million cell updates/sec

Title: US-09-546-136-5

Perfect score: 723

Sequence: 1 MELKQAFVPEFENLSSSG.....LNKMLLSILLGNEENLLQIS 144

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109.5	15.1	902	1 RNE_BUCAI	P57429 buchnera ap
2	98.5	13.6	1044	1 VAF3_SCHPO	Q09857 schizosacch
3	95.5	13.2	650	1 MTLR_STRMU	Q02425 streptococc
4	95	13.1	404	1 ALR_RICPR	Q92e52 rickettsia
5	92	12.7	602	1 PRIM_MYCPU	Q98qb3 mycoplasma
6	92	12.7	873	1 SYA_WIGBR	Q8d2w8 wiggleswort
7	91	12.6	377	1 YPJH_BACSU	P42982 bacillus su
8	90.5	12.5	297	1 RGG_STRGC	P49330 streptococc
9	90.5	12.5	453	1 ENGA_BUCAP	O51881 buchnera ap
10	89.5	12.4	799	1 SYL_MYCPE	Q8ew18 mycoplasma
11	87.5	12.1	159	1 OBP_BOVIN	P07435 bos taurus
12	87.5	12.1	386	1 RMAR_WANNI	P48849 hansenula w
13	87	12.0	355	1 ALR_RICCN	Q92jd9 rickettsia
14	87	12.0	575	1 RPOC_ASTLO	P58131 astasia lon
15	87	12.0	811	1 RPOP_NEUIN	P33541 neurospora
16	86.5	12.0	620	1 Y241_MYCJE	P47483 mycoplasma
17	86	11.9	198	1 ENGB_CAMGE	Q9ph17 campylobact
18	86	11.9	206	1 LOLA_WIGBR	Q8d264 wiggleswort
19	86	11.9	442	1 TIG_BUCAI	P57546 buchnera ap
20	85.5	11.8	380	1 RMAR_WILMR	P47906 willipopsis
21	85.5	11.8	1155	1 IF2P_METJA	Q57710 methanococc
22	85	11.8	414	1 YR28_BORBU	O50999 borrelia bu
23	85	11.8	814	1 AKH_BUCAP	Q8K9u9 buchnera ap
24	84.5	11.7	575	1 RPOC_PLAFA	P21422 plasmodium
25	84.5	11.7	609	1 YSW1_YEAST	P38280 saccharomyc
26	84.5	11.7	1121	1 Y309_YEAST	P47107 saccharomyc
27	84	11.6	861	1 SYL_BUCAP	Q8K9D9 buchnera ap
28	83.5	11.5	368	1 TRMU_BUCAI	P57349 buchnera ap
29	83.5	11.5	378	1 PYRX_HELPJ	Q92mg9 helicobacte
30	83.5	11.5	481	1 SYEL_THETN	Q8tb93 thermomanae
31	83	11.5	212	1 DSBA_BUCAI	P57505 buchnera ap
32	83	11.5	862	1 MUTS_BORBU	O51737 borrelia bu
33	82.5	11.4	339	1 RMAR_CANGA	P21358 candida gla

34 82.5 11.4 503 1 SYK_BUCAP
35 82.5 11.4 1076 1 RPOB_ASTLO
36 82 11.3 336 1 YD48_METJA
37 82 11.3 453 1 TRME_WIGBR
38 82 11.3 474 1 Y381_BORBU
39 82 11.3 569 1 CYSP_PLAFA
40 81.5 11.3 490 1 PIT_BUCAI
41 81.5 11.3 1041 1 TLRB_HUMAN
42 81.5 11.3 1155 1 KOMI_YEAST
43 81 11.2 598 1 TDPI_CAEEL
44 81 11.2 628 1 GIDA_WIGBR
45 81 11.2 1636 1 YN37_YEAST

ALIGNMENTS

RESULT 1
RNE_BUCAI
ID RNE_BUCAI STANDARD; PRT: 902 AA.
AC P57429;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Ribonuclease E (EC 3.1.4.-) (RNase E).
GN RNE OR BU347.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
RP [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
Buchnera sp. APS.";
RL Mature 407:81-86(2000).
CC -!- FUNCTION: THIS PROTEIN MATURES 5S RRNA FROM ITS PRECURSORS FROM
ALL THE RRNA GENES. IT IS THE MAJOR ENDOBIONUCLEASE PARTICIPATING
IN MRNA TURNOVER (BY SIMILARITY).
CC -!- SUBUNIT: ORGANIZED INTO A STRUCTURE (PROCESSOME OR RNA
DEGRADOSOME) CONTAINING A NUMBER OF RNA-PROCESSING ENZYMES (BY
SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE RNE FAMILY.
CC -!- SIMILARITY: Contains 1 SI motif domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AP001119; BAB13052.1; -
CC InterPro: IPR004659; RNaseEG.
CC InterPro: IPR003029; SI.
CC Pfam: PF00575; SI; 1.
CC SMART: SM00316; SI; 1.
CC TIGRfams: TIGR00757; RNaseEG; 1.
CC PROSITE: PS50126; SI; 1.
KW Hydrolase; Nuclease; Endonuclease; RNA-binding; Complete proteome.
FT DOMAIN 39 119 SI MOTIF.
SQ SEQUENCE 902 AA; 104006 MW; C3FA00476E3C2E37 CRC64;
Query Match 15.1%; Score 109.5; DB 1; Length 902;
Best Local Similarity 28.6%; Pred. No. 0.25;
Matches 44; Conservative 26; Mismatches 55; Indels 29; Gaps 6;
QY 10 EFDEN-LSSSGSIHLKVKQNCSPNYDFKITFDIGLYIKNK--SGVLDKYLKNNV 66

```

Db 544 EAQENILKNNYNILKVLGN-NRNFIFKMITFKNSFFIKNNMLTSDIFKKNLTKNT- 601
Oy 67 SLVAKRDYLSLSNNKQIKKF-----KNIKNKHKN-----KFN 102
Db 602 NIFFKKYSLNNKNNQKRVLSKLFENIENIPKNNKLDLTSSANYLYDNIERKKN 661
Oy 103 LYVNEDIKRTKNGILEEVILKNNMLLSILLGN 136
Db 662 ITRKNDLIQNIHENSYLKHLVMNRYNVIINN 695

RESULT 2
YAF3_SCHPO
ID YAF3_SCHPO STANDARD; PRT; 1044 AA.
AC Q09857; Q097V0;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C29E6.03c in chromosome I.
GN SPAC29E6.03c OR SPAC30.07C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN RP SEQUENCE FROM N.A.
RC MEDLINE=21848401; PubMed=118593360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Spouris J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds R., Squares R., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Art R., Robben J., Grymonprez B.,
RA Weltjens I., vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shipakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RL "The genome sequence of Schizosaccharomyces pombe."
CC CC -!- SIMILARITY: BELONGS TO THE VDP/USOI/YBLO47C FAMILY.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC
CC DR EMBL: Z66525; CAA91425.1; ALT_SEQ.
CC DR PIR: A133538; CAB6466.1;
CC DR GeneDB: SPAC29E6.03c;
CC DR InterPro: IPR006953; Usol_p115.head.
CC DR Pfam: PF04869; Usol_p115.head; 1
CC DR Hypothetical protein; Coiled coil.

```

```

FT DOMAIN 600 1014 COILED COIL (POTENTIAL).
SQ SEQUENCE 1044 AA; 119127 MW; 52ACE0E174725A59 CRC64;

Query Match 13.6%; Score 98.5; DB 1; Length 1044;
Best Local Similarity 27.9%; Pred. No. 2.1;
Matches 39; Conservative 27; Mismatches 43; Indels 31; Gaps 8;

Oy 22 IHLEKVKONSPNYDFKTFIDCYL-----YIKNKGSVILDKYDLKKNVLSLALK-RDY 75
Db 673 IELDYATSNCKQMEEMQV-LRGHESEIKDFIEHSEKLTQKDDIKQFGIILSKNRDL 731
Oy 76 LSLSLSNKQIKFKFNI-----KNHKLKNFNLVINEDIKRT-----KNGILE 121
Db 732 LS-----ELEKSKLSNLSLALESKNKLENDNL--LTKLNKKNADTESFKNTI-R 781
Oy 122 EVILNKMLLSILLGNENLL 141
Db 782 EAELSKKALNDNLGNENII 801

RESULT 3
MTLR_STRMU
ID MTLR_STRMU STANDARD; PRT; 650 AA.
AC Q02425;
DT 01-JUL-1993 (Rel. 26, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative transcriptional regulator mtlr.
GN MTLR OR SMU.1184C.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN RP SEQUENCE FROM N.A., AND FUNCTION.
RC STRAIN=UA130 / Serotype C;
RX MEDLINE=20340960; PubMed=10878121;
RA Honeyman A.L., Curtiss R. III;
RT "The mannitol-specific enzyme II (mtIIA) gene and the mtlr gene of the
RL Pts of Streptococcus mutans."
RL Microbiology 146:1565-1572(2000).
RN RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RL pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RN RP SEQUENCE OF 525-650 FROM N.A.
RC STRAIN=UA130 / Serotype C;
RX MEDLINE=92348013; PubMed=1322373;
RA Honeyman A.L., Curtiss R. III;
RT "Isolation, characterization, and nucleotide sequence of the
RL mannitol-specific factor III gene of the phosphoenolpyruvate
RL phosphotransferase system."
RL Infect. Immun. 60:3369-3375(1992).
CC CC -!- FUNCTION: NOT NECESSARY FOR MANNITOL UTILIZATION. MAY BE INVOLVED
CC CC IN REGULATION OF THE MANNITOL PHOSPHOENOLPYRUVATE-DEPENDENT SUGAR
CC CC PHOSPHOTRANSFERASE SYSTEM (PTS).
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC

```


DR EMBL; AF210133; AAA26940.2; -
 DR EMBL; AE014955; AAN58874.1; -
 DR PIR; A44798; A44798.
 KW Transcription regulation; Complete proteome.
 FT CONFLICT 361 363 KEP -> RES (IN REF. 1).
 FT CONFLICT 441 441 V -> A (IN REF. 1).
 FT CONFLICT 447 447 M -> I (IN REF. 1).
 FT CONFLICT 482 482 K -> R (IN REF. 1).
 FT CONFLICT 539 539 C -> R (IN REF. 1 AND 3).
 FT CONFLICT 551 551 H -> R (IN REF. 1 AND 3).
 FT CONFLICT 620 620 L -> P (IN REF. 1 AND 3).
 FT CONFLICT 620 620 L -> P (IN REF. 1 AND 3).
 SQ SEQUENCE 650 AA; 75337 MW; 02F1229435B7C97A CRC64;
 Query Match 13.28; Score 95.5; DB 1; Length 650;
 Best Local Similarity 25.88; Pred. No. 2.2;
 Matches 42; Conservative 30; Mismatches 44; Indels 47; Gaps 9;
 QY 1 MELKQAFVFEEDENLSSSGSIHLEKVKQNCSPNYDYFKITFDGVLKYNKSGVILDKY 60
 Db 29 LKVSQRTVYRTISDLTDSLTNTISIKEN--QNY-----FLVG----- 65
 QY 61 DLKNVISLALK--RDYLSLSLNKKQIKKFKNKNKLNKFNLYVIN-----EDIEKR 113
 Db 66 ELANLASIISLDYEQYERLNLITYKLMSFSITHEQLOEERFNVSNVTIIQIDAEIEKR 125
 QY 114 IT-----KNGILEVILNK-----MLJSILLGNEENLQIS 144
 Db 126 LADFOLRLDRKGY--RLVGKNTLRLLAILLTNN---LSIS 163
 RESULT 4
 ALR_RICPR STANDARD: PRT: 404 AA.
 AC Q9Z52;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Alanine racemase (EC 5.1.1.1).
 GN ALR OR RP095.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsia.
 OX NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madrid E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 RT mitochondria."
 RL Nature 396:133-140(1998).
 RN [2]
 RP DOMAIN RPEL.
 RX MEDLINE=20485642; PubMed=11030655;
 RA Ogata H., Audic S., Barbe V., Artiguenave F., Fournier P.-E.,
 RA Raoult D., Claverie J.-M.;
 RT "Selfish DNA in protein-coding genes of Rickettsia."
 RL Science 290:347-350(2000).
 CC -!- FUNCTION: Provides the D-alanine required for cell wall
 CC biosynthesis (By similarity).
 CC -!- CATALYTIC ACTIVITY: L-alanine = D-alanine.
 CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
 CC -!- PATHWAY: D-alanine branch of peptidoglycan biosynthesis; first
 CC step.
 CC -!- SIMILARITY: Belongs to the alanine racemase family.
 CC -!- SIMILARITY: Contains 1 RPEL insert domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ235270; CAA14565.1; -
 DR PIR; F71718; F71718.
 DR HAMAP; MF_01201; atypical; 1.
 DR InterPro; IPR000821; Ala_racemase.
 DR InterPro; IPR005728; Rickett_RPE.
 DR Pfam; PF00842; Ala_racemase_C; 1.
 DR Pfam; PF01168; Ala_racemase_N; 1.
 DR PRINTS; PR00992; ALARACEMASE.
 DR TIGRFAMS; TIGR01045; RPE; 1.
 DR PROSITE; PS00395; ALANINE_RACEMASE; 1.
 KW Isomerase: Pyridoxal phosphate; Cell wall; Peptidoglycan synthesis;
 KW Complete proteome. 273
 FT DOMAIN 226 34 CATALYTIC BASE SPECIFIC TO D-ALANINE (BY
 FT ACT_SITE 34 34 CATALYTIC).
 FT ACT_SITE 298 298 CATALYTIC BASE SPECIFIC TO L-ALANINE (BY
 FT ACT_SITE 34 34 CATALYTIC).
 FT BINDING 34 34 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 FT SEQUENCE 404 AA; 44730 MW; FEA23D86698256DA CRC64;
 SQ
 Query Match 13.18; Score 95; DB 1; Length 404;
 Best Local Similarity 27.08; Pred. No. 1.5;
 Matches 41; Conservative 27; Mismatches 44; Indels 40; Gaps 10;
 QY 20 GSIHLEK--VKQNCSPNYDYFKITFDGVLKYNKSGVILDKYNVISLALK---R 73
 Db 41 GAOIAKALIKENC---QYFFVA-----TSEGINLRKV-LNNDITILVINGVFT 87
 QY 74 DYLSL-----SLSNKKQI---KKFKNKNK-----HLKNFNLVINEED-LEKRITK 116
 Db 88 DALEIQTNYLTPVNLNLSQIEIQKFSNLKGLKILPCVLFHNTGLNRLGNSDEIQLEQLND 147
 QY 117 NGILE----EVILNKMLLSILLGNEENLQIS 144
 Db 148 RDLKGLDLQYIISHLAASEETGNPNLIQLN 179
 RESULT 5
 PRIM_MYCPU STANDARD: PRT: 602 AA.
 AC Q9BOB3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA primase (EC 2.7.7.-).
 GN DNAG OR MYPV_4530.
 OS Mycoplasma pulmonis.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UAB CTIP;
 RX MEDLINE=21267165; PubMed=11353084;
 RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
 RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
 RA Blanchard A.;
 RT "The complete genome sequence of the murine respiratory pathogen
 RT Mycoplasma pulmonis."
 RL Nucleic Acids Res. 29:2145-2153(2001).
 CC -!- FUNCTION: DNA PRIMASE IS THE POLYMERASE THAT SYNTHESIZES SMALL
 CC RNA PRIMERS FOR THE OKAZAKI FRAGMENTS ON BOTH TEMPLATE STRANDS AT
 CC REPLICATION FORKS DURING CHROMOSOMAL DNA SYNTHESIS.
 CC -!- COFACTOR: BINDS ONE ZINC ION PER MOLECULE (BY SIMILARITY).
 CC -!- SUBUNIT: Monomer (By similarity).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AL445564; CAC13626.1; -
DR PIR; E90568; E90568.
DR MYPULIST; MYPU_4530; -
DR InterPro; IPR006295; DNA_primase.
DR InterPro; IPR006171; Toprim.dom.
DR InterPro; IPR006647; Toprim_primase.
DR InterPro; IPR006154; Toprim_sub.
DR InterPro; IPR002694; Znf_CHC2.
DR Pfam; PF01751; Toprim; 1.
DR Pfam; PF01807; zf-CHC2; 1.
DR ProDom; PD002276; Toprim_primase; 1.
DR ProDom; PD002988; Znf_CHC2; 1.
DR SMART; SM00493; TOPRIM; 1.
DR TIGRFAMS; TIGR01391; dnaG; 1.
KW Transferase; DNA replication; DNA-directed RNA polymerase; Primosome;
KW Zinc-finger; Zinc; Metal-binding; Complete proteome.
FT ZN_FING 38 62 CHC2-TYPE (BY SIMILARITY).
SQ SEQUENCE 602 AA: 70907 MW: 01FIC333687049BD CRC64;

Query Match 12.7%; Score 92; DB 1; Length 602;
Best Local Similarity 26.7%; Pred. No. 3.8;
Matches 40; Conservative 24; Mismatches 48; Indels 38; Gaps 8;
QY 12 DENLSSSGSIHLEKVKQNCSPNYDFKTFIDGVIYIKNKGVLIDKYNVLSVAL 71
DB 338 DELLKOKGANFLKELIEKKNKPAID-----YLY-----EFLTSKYNNINNSEHLISF 384
QY 72 KRDYLS-LSLSNKKQIKKFNKKNHLKN--KFNLYV-----INE--DIEKRI----- 114
DB 385 VNDPAPLLSSQNNVIDKYENLLMKHSINIRFLYONKQYDRINOYNSLQORFEEPKQ 444
QY 115 -----TKNGILEVILN-----KMLLSILLGN 136
DB 445 KKKANSREELKQYONLYVKLLLSLNSN 474

RESULT 6

SYA_WIGBR STANDARD; PRT; 873 AA.
AC Q8D2W8;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (AlaRS).
GN ALAS OR WIGBR2340.
OS Wigglesworthia glossinidia brevipalpis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Wigglesworthia.
OX NCBI_TaxID=36870;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2297718; PubMed=12219091;
RA Akman L., Yamashita A., Watanabe H., Oshima K., Shiba T., Hattori M.,
RA Aksoy S.;
RT "Genome sequence of the endocellular obligate symbiont of tsetse
RT flies, Wigglesworthia glossinidia".
RL Nat. Genet. 32:402-407(2002).
CC -|- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
CC diphosphate + L-alanyl-tRNA(Ala).
CC -|- SUBUNIT: Homotrimer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; AB063521; BAC24380.1; -
DR HAMAP; MF_00036; -; 1.
DR InterPro; IPR002318; tRNA-synt_2c.
DR InterPro; IPR006193; tRNA-synt_Ala.
DR Pfam; PF01411; tRNA-synt_2c; 1.
DR PRINTS; PR00980; TRNASYNTHALA.
DR TIGRFAMS; TIGR00344; alas; 1.
DR PROSITE; PS00860; AA-TRNA-LIGASE_II_ALA; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 873 AA: 101257 MW: 6EC025D0A1E6A9EC CRC64;

Query Match 12.7%; Score 92; DB 1; Length 873;
Best Local Similarity 27.5%; Pred. No. 5.7;
Matches 41; Conservative 15; Mismatches 43; Indels 50; Gaps 5;
QY 45 GYLTKNKGVLIDKYNVLSVALK-----RDYLSLSLNKKQI----- 86
DB 673 GCFVITNFSKISSDYRIKAITGETAFAIOKKFENDINTISLLIKSNSNEIVNKIEKITE 732
QY 87 -----KKFNKKNKHLKN-----KFNLYV-----NEDIEKRI----- 114
DB 733 NLKILEKKNKLNKKNKIKKLNFKIKKINFEITMDYFNEEDIKIILREILNKVKTN 792
QY 115 TKNGILEVILNKKMLLSILLGNENLLOI 143
DB 793 LKNGVILASIKKKNLYITVGTNNLVNI 821

RESULT 7

YPJH_BACSU
ID YPJH_BACSU STANDARD; PRT; 377 AA.
AC P42982;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative glycosyl transferase ypjH (EC 2.-.-.-).
GN YPJH OR JOJH.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / Marburg;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RA Serron P.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between
RT the serA and kds loci cloned in a yeast artificial chromosome.";
RL Microbiology 142:2005-2016(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Frits C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gim S.-Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portetelle D., Porvollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat R., Yamamoto H., Yamane K., Yasumoto K., Yata A.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
 RT *The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*.*
 RL Nature 390:249-256(1997).
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L38424; AAA92877.1; ALT_SEQ.
 DR EMBL: L47709; AAB38445.1; -.
 DR EMBL: Z99115; CAB14162.1; -.
 DR PIR: G69937; G69937.
 DR Subtilist: BG11213; yvjH.
 DR InterPro: IPR001296; Glyco.trans_1.
 DR Pfam: PF00534; Glycos_transf_1; 1.
 DR KW Hypothetical protein; Transferase; Glycosyltransferase;
 KW Complete proteome.
 SQ SEQUENCE 377 AA; 41978 MW; 2DB1AB344D6536AA CRC64;

Query Match 12.6%; Score 91; DB 1; Length 377;
 Best Local Similarity 23.0%; Pred. No. 2.8;
 Matches 37; Conservative 31; Mismatches 43; Indels 50; Gaps 8;
 QY 7 FVFEEDENLSSSGSIHLEKVKQNCSPNYDYFK-----ITFDGYLYIKNKGSGVILD 58
 DB 140 FAIESSDRVTAVSSALAE-----TYDLIKPEKTIETIYFNFDVRLKKNNTAAIKE 191
 QY 59 KYDLKNVLSLVALKRDYLSLSNNKQTK-----PKNTKNKHLKKNFNLVINEDIE 111
 DB 192 KHGI-----LPDEKVIHVSNPKRVQDVIRVFRNTAG---KTKAKLLLVGDGPE 240
 QY 112 KR-----ITKNCILEEVILNKLKLLSILLGNE---ENLLQIS 144
 DB 241 KSTACELIRKYGLEQDV-----LMLGNODRVEDLYSIS 273

RESULT 8
 RCG_STRCC
 ID RCG_STRCC STANDARD; PRT; 297 AA.
 AC P49330;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Protein r9g.
 GN R9G.
 OS Streptococcus gordonii Challis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=29390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92276337; PubMed=1534326;
 RA Sulavik M.C., Tardif G., Clewell D.B.;
 RT "Identification of a gene, r9g, which regulates expression of
 RT glycosyltransferase and influences the Spp phenotype of *Streptococcus*
 RT *gordonii* Challis.";
 RL J. Bacteriol. 174:3577-3586(1992).

CC -!- FUNCTION: REGULATES THE EXPRESSION OF GLYCOSYLTRANSFERASE (GTF6).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (potential).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M89776; AAA26968.1; -.
 DR PIR: A41898; A41898.
 DR InterPro: IPR001387; HTH_3.
 DR Pfam: PF01381; HTH_3; 1.
 DR SMART: SM00530; HTH_XRE; 1.
 KW Transcription regulation; DNA-binding.
 SQ SEQUENCE 297 AA; 34439 MW; A24D3BB3C67FF5E CRC64;
 Query Match 12.5%; Score 90.5; DB 1; Length 297;
 Best Local Similarity 24.5%; Pred. No. 2.4;
 Matches 39; Conservative 31; Mismatches 54; Indels 35; Gaps 8;
 QY 5 QAFVFEEDENLSSSGSIHLEKVKQNCSPNYDYFK--ITFDGYLYIKNKGSGVILDKYDL 62
 DB 103 EAMAEQEPPEKKYKLTIVIRATLTSNPDYQVSKGDIEFLDYLSFVSEWG----RYEL 158
 QY 63 ---KNVISLVALK-----RDYLSLSNNKQIKKFN-----LKNKHLK--N 99
 DB 159 WLFNSVNLLETLETLETFASEMINRTQYNNLPENRRRIKMLNIVSVSACIENHLLQVAM 218
 QY 100 KFNLYVINEDIKRIKNGILEEVIL--NKMLLSILLGN 136
 DB 219 KELNYIDN---TKIPETDLYRVLKLYHKALYSYKVG 253

RESULT 9
 ENGA_BUCAP
 ID ENGA_BUCAP STANDARD; PRT; 453 AA.
 AC O51881;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable GTP-binding protein *enga*.
 GN ENGA OR BUSG382.
 OS Buchnera aphidicola (subsp. *Schizaphis graminum*).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Buchnera.
 OX NCBI_TaxID=98794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98184963; PubMed=9516544;
 RA Clark M.A., Baumann L., Baumann P.;
 RT "Sequence analysis of a 34.7-kb DNA segment from the genome of
 RT Buchnera aphidicola (endosymbiont of aphids) containing *groEL*, *dnaA*,
 RT the *atp* operon, *gidA*, and *rho*.";
 RL Curr. Microbiol. 36:158-163(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22084549; PubMed=12089438;
 RA Tamas I., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
 RA Wernegreen J.J., Sandstrom J.P., Moran N.A., Andersson S.G.E.;
 RT "50 million years of genomic stasis in endosymbiotic bacteria.";
 RL Science 296:2376-2379(2002).
 CC -!- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING
 CC PROTEINS. ENGA SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

or send an email to license@isb-sib.ch.

```
CC CC
CC EMBL; AF008210; AAC38119.1;
CC DR EMBL; AE014132; AAM68116.1;
CC DR HAMAP; MF_00195; --; 1.
CC DR InterPro; IPR005289; GTP-binding_dom.
CC DR InterPro; IPR006073; GTP_OBG.
CC DR InterPro; IPR002917; MMR_HSR1.
CC DR InterPro; IPR001806; Ras_transfmrng.
CC DR InterPro; IPR005225; Small_GTP.
CC DR Pfam; PF01926; MMR_HSR1.1.
CC DR PRINTS; PR00326; GTP_OBG.
CC DR PRINTS; PR00449; RASTRNSFRMG.
CC DR TIGRFAMS; TIGR00650; MG442; 2.
CC DR TIGRFAMS; TIGR00231; small_GTP; 2.
CC KW GTP-binding; Repeat; Complete proteome.
CC FT NP_BIND 9 16 GTP 1 (POTENTIAL).
CC FT NP_BIND 57 61 GTP 1 (POTENTIAL).
CC FT NP_BIND 119 122 GTP 1 (POTENTIAL).
CC FT NP_BIND 194 201 GTP 2 (POTENTIAL).
CC FT NP_BIND 241 245 GTP 2 (POTENTIAL).
CC FT NP_BIND 306 309 GTP 2 (POTENTIAL).
CC SQ SEQUENCE 453 AA; 52047 MW; 8B7A3BC29DD53B42 CRC64;

Query Match 12.5%; Score 90.5; DB 1; Length 453;
Best Local Similarity 27.6%; Pred. No. 3.7;
Matches 43; Conservative 23; Mismatches 59; Indels 31; Gaps 7;

QY 13 ENLSSSSGSIHLEKVKQNSPNVDYFKITFDIGLYIK-----NKSGVILD-----KYD 61
DB 174 EKISQHSKKNKNSVIAICIGPNVGKSTLINSLLMKKRMITSNKAGTTLDTVLVPIKNY 233
QY 62 LKNVISLVALKRDYLSLSLSNNK--OIKKFKIKKHLKKNFLNLYVINEIDIEKRIKNGI 119
DB 234 YKNYIFI-----DTAGSKSKKSKTKNIEKFKIKTLQTIKESHLTLILIDAKDQISKQDL 288
QY 120 LEE-----VILNKM-LISIL-LGNEENLQ 142
DB 289 LLSSFTEKSGKPLIIVINKCDLLSLKEKNLENLIK 324

RESULT 10
SYL_MYCPE
ID SYL_MYCPE STANDARD; PRT; 799 AA.
AC Q8EW18;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Leucyl-tRNA synthetase (EC 6.1.1.4) (Leucine--tRNA ligase) (Leurs).
GN LEUS OR MYPE3890.
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=28227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HF-2;
RX MEDLINE=22354719; PubMed=12466555;
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans.";
RL Nucleic Acids Res. 30:5293-5300(2002).
CC -!- CATALYTIC ACTIVITY: ATP + L-leucine + tRNA(Leu) = AMP +
CC diphosphate + L-leucyl-tRNA(Leu).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
```

or send an email to license@isb-sib.ch.

```
CC CC
CC EMBL; AP004171; BAC44178.1;
CC DR HAMAP; MF_00049; --; 1.
CC DR TIGRFAMS; TIGR00396; leus_bact; 1.
CC DR PROSITE; PS00178; AA_TRNA_LIGASE_I.1.
CC KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC FT SITE 39 50 "HIGH" REGION.
CC FT SITE 575 579 "KMSKS" REGION.
CC FT BINDING 578 578 ATP (BY SIMILARITY).
CC SQ SEQUENCE 799 AA; 93329 MW; 33B62E3CD73D06FE CRC64;

Query Match 12.4%; Score 89.5; DB 1; Length 799;
Best Local Similarity 26.1%; Pred. No. 8.1;
Matches 36; Conservative 21; Mismatches 42; Indels 39; Gaps 5;

QY 7 FVFEDENLSSSSGSIHLEKVKQNSPNVDYFK-----ITFDIGLYIKNK-- 52
DB 661 FNEEDFNLVISEMIFINECYKYEKINYDMLNFCVILSCFAPFITEINEVFLKNKF 720
QY 53 -SGVLDKYDLKNV-----ISLVALKRDYLSLSLSNNKQIKKFN 91
DB 721 ISDNLMPKYDEKKIVETIKIPVQINGKIREVLEINLGATQKDVVDLAIKNEKIWKIEN 780
QY 92 ---IKNHLKKN-FNLV 105
DB 781 KKVKEIYIENKILNLI 798

RESULT 11
OBP_BOVIN
ID OBP_BOVIN STANDARD; PRT; 159 AA.
AC P07435;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Odorant-binding protein (OBP) (Olfactory mucosa pyrazine-binding
DE protein).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=90076175; PubMed=2512125;
RA Tirindelli R., Keen J.N., Cavaggioni A., Eliopoulos E.E.,
RA Findlay J.B.C.;
RT "Complete amino acid sequence of pyrazine-binding protein from cow
RT nasal mucosa.";
RL Eur. J. Biochem. 185:569-572(1989).
RN [2]
RP SEQUENCE OF 1-63.
RX MEDLINE=87134269; PubMed=3817156;
RA Cavaggioni A., Sorbi R.T., Keen J.N., Pappin D.J.C., Findlay J.B.C.;
RT "Homology between the pyrazine-binding protein from nasal mucosa and
RT major urinary proteins.";
RL FEBS Lett. 212:225-228(1987).
RN [3]
RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=92322903; PubMed=1623143;
RA Monaco H.L., Zanotti G.;
RT "Three-dimensional structure and active site of three hydrophobic
RT molecule-binding proteins with significant amino acid sequence
RT similarity.";
RL Biopolymers 32:457-465(1992).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=97057532; PubMed=8901871;
RA Bianchet M.A., Bains G., Pelosi P., Pevsner J., Snyder S.H.,
RA Monaco H.L., Amzel L.M.;
```

OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC	Saccharomycetales; Saccharomycetaceae; Pichia.
OC	NCBI_TaxID=4907;
OX	[1]
RN	SEQUENCE FROM N.A.
RP	STRAIN=21;
RC	Sekito T., Okamoto K., Kitano H., Yoshida K.;
RA	"Yeast Hansenula wingii mitochondria genome's complete DNA sequence
RT	demonstrated unique characteristics.";
RL	Nucleic Acids Symp. Ser. 31:233-234(1994).
CC	-!- FUNCTION: ESSENTIAL FOR MITOCHONDRIAL PROTEIN SYNTHESIS AND
CC	REQUIRED FOR THE MATURATION OF SMALL RIBOSOMAL SUBUNITS
CC	(BY SIMILARITY).

```

CC      1-1 SUBCELLULAR LOCATION: Mitochondrial.
CC      1-1 SIMILARITY: BELONGS TO THE VAR1 FAMILY OF RIBOSOMAL PROTEINS.
CC      -----
CC      CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      CC the European Bioinformatics Institute. There are no restrictions on its
CC      CC use by non-profit institutions as long as its content is in no way
CC      CC modified and this statement is not removed. Usage by and for commercial
CC      CC entities requires a license agreement (See http://www.isb-sib.ch/announcements
CC      CC or send an email to license@isb-sib.ch).
CC      -----
CC      CC EMBL; D31785; BAA06578.2; -.
CC      CC DR PIR; S58755; S58755.
CC      CC DR Pfam; PF05316; Yeast_VAR1.1.
CC      CC KW Ribosomal protein; Mitochondrion.
CC      CC SQ SEQUENCE 386 AA; 44904 MW; 9AD0D625F303E88F CRC64;
CC
Query Match 12.1%; Score 87.5; DB 1; Length 386;
Best Local Similarity 33.3%; Pred. No. 5.4;
Matches 39; Conservative 20; Mismatches 39; Indels 19; Gaps 7
QY 39 KITFDIGLYLIKNGSGV-----ILDKYOLKNVISLVALKRDYLSLSLSNKKQIK----- 87
Db 44 KINSWDNQLYKFNKNVINTWILDRDYSKLKIKFKVVRINNNINNGOIKDYINKP 103
QY 88 KFNKIKNKLKKNFLNLYVINEDIKRTKNGCILEEVLNKKMLLSILLGNEENLLOIS 144
Db 104 KFKHTINKVYN-FN-YILSSN---NITINDI--DNNKNKYTST-INDINNLOIS 152

```

```

ALR_RICCN          STANDARD;          PRT;      355 AA.
AC Q92JD9;
AC Q92JD9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alanine racemase (EC 5.1.1.1).
DE ALR OR RC0128.
GN Rickettsia conorii.
OS Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OC NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Sanson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
RC -!- FUNCTION: Provides the D-alanine required for cell wall
CC biosynthesis (By similarity).
CC -!- CATALYTIC ACTIVITY: L-alanine = D-alanine.
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC -!- PATHWAY: D-alanine branch of peptidoglycan biosynthesis; first
CC step.
CC -!- SIMILARITY: Belongs to the alanine racemase family.
CC -----

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/or_send_an_email_to_license@sib-sib.ch).

CC
CC
On send an email to license@rsb.sib.ch).

EMBL; AJ294725; CAC24618.1; -;
DR DR HSSP; O9KWU6; 1HOM.
DR DR InterPro: IPR000722; RNA_pol_A.
DR DR InterPro: IPR007080; RNA_pol_Rpb1_1.
DR DR InterPro: IPR006592; RNA_polA_N.
DR DR Pfam: PF04093; RNA_pol_Rpb1_1; 1.
DR DR Pfam: PF00623; RNA_pol_Rpb1_2; 1.
DR DR SMART; SM00663; RPOLA_N; 1.
DR DR Transference; Transcription; DNA-directed RNA polymerase; Chloroplast.
SQ SEQUENCE 575 AA; 67075 MW; C83269DOB54F8DAF CRC64;

	Query Match	12.0%	Score 87;	DB 1;	Length 575;
	Best Local Similarity	26.7%;	Pred. No. 9;		
	Matches	46;	Conservative 24;	Mismatches 54;	Indels 48; Gaps 9;
Oy	13 ENLSSSGSHLEKVKONCSPNYDYKPTIFDGLYIKKSGVILDKYLKNVISLVAK	72			
Db	161 EAYNDILSKIDLEKTUUKTNTHQYK---EKNTIKNFKN-ILANKILYNLLQTKIK	215			
Oy	73 RQYL-----SLSLNNKOIK-----KPKNIKHHLKHNENL----	YVI 106			
Db	216 FSWLLKFYLPVPNPYPPIINMKNNQOISNDNLTYASIIINVANKIKLESIPONYFI	275			
Oy	107 NEDI--EKRIT-----KNGILEEVILNKMLLSI----DLGNE----	ENLL 141			
Db	276 NEKILLQKKVDOLINNEKYKENLGKIINNKKLSITENIKKGEGIIRENML	327			

```

RESULT 15
RPOP_NEUIN
ID RPOP_NEUIN STANDARD; PRT; 811 AA.
AC P33541;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable DNA-directed RNA polymerase (EC 2.7.7.6).
OS Neurospora intermedia.
OG Mitochondrion.
OG Plasmid kalilo.
OG Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5142;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=92035090; PubMed=1934129;
RA Chan B.S.S., Court D.A., Vierula P.J., Bertrand H.;
RT "The kalilo linear senescence-inducing plasmid of Neurospora is an
RT invertion and encodes DNA and RNA polymerases.";
RL Curr. Genet. 20:225-237(1991).
RN [2]
RN SEQUENCE OF 1-58 FROM N.A.
RX MEDLINE=93024309; PubMed=1406582;
RA Vierula P.J., Bertrand H.;
RT "A deletion derivative of the kalilo senescence plasmid forms hairpin
RT and duplex DNA structures in the mitochondria of Neurospora.";
RL Mol. Gen. Genet. 234:361-368(1992).
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC {RNA}(N).
CC -!- SIMILARITY: BELONGS TO THE PHAGE AND MITOCHONDRIAL RNA POLYMERASES
CC FAMILY.
CC

```

	EMBL; X52106;	CAG36326.1;	-	
CC	DR	PIR; S17908;	S17908.	
DD	DR	HSP; P00573;	IARO.	
DE	DR	InterPro:	IPRO02092;	RNA_pol_phase.
DF	DR	Pfam;	PF00940;	RNA_pol_1.
DG	DR	PROSITE;	PSO0489;	RNA_POL_PHASE_2; 1.
DH	DR	PROSITE;	PSO0900;	RNA_POL_PHASE_1; 1.
DI	KW	Transferase;	Transcription;	DNA-directed RNA polymerase;
DJ	KW	Mitochondrion;	Plasmid.	
EI	FT	ACT_SITE	470-470	BY SIMILARITY.
EL	FT	ACT_SITE	540-540	BY SIMILARITY.
EQ	FT	ACT_SITE	718-718	BY SIMILARITY.
ER	SO	SEQUENCE	811 AA;	94800 MW; 286225A36B625BBC CRC64;

Search completed: July 24, 2003, 20:06:34
Job time : 27.2468 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:01:59 ; Search time 109.403 Seconds

(without alignments)
339.659 Million cell updates/sec

Title: US-09-546-136-5

Perfect score: 723

Sequence: 1 MELQAFVFEDENLSSSG.....LNKMLLSILLGNEENLLOIS 144

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*

- 1: sp-archaea.*
- 2: sp-bacteria.*
- 3: sp-fungi.*
- 4: sp-human.*
- 5: sp-invertebrate.*
- 6: sp-mammal.*
- 7: sp-mhc.*
- 8: sp-organelle.*
- 9: sp-phage.*
- 10: sp-plant.*
- 11: sp-rodent.*
- 12: sp-virus.*
- 13: sp-vertebrate.*
- 14: sp-unclassified.*
- 15: sp-virus.*
- 16: sp-bacteriap.*
- 17: sp-archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	723	100.0	144	2 052975	O52975 clostridium
2	553	76.5	142	2 050600	O50600 clostridium
3	540	74.7	142	2 050597	O50597 clostridium
4	111.5	15.4	2279	5 081JP6	O81JP6 plasmodium
5	110.5	15.3	1782	5 081L21	O81L21 plasmodium
6	106	14.7	796	5 081EN5	O81EN5 plasmodium
7	105.5	14.6	607	5 081CC9	O81CC9 plasmodium
8	105.5	14.6	1213	5 081C20	O81C20 plasmodium
9	105.5	14.6	2961	5 081S75	O81S75 plasmodium
10	104.5	14.5	1001	16 025314	O25314 helicobacte
11	104.5	14.5	1226	5 081L77	O81L77 plasmodium
12	104	14.4	245	17 081JRO	O81JRO methanosarc
13	102.5	14.2	227	5 081CK6	O81CK6 plasmodium
14	102	14.1	4202	5 081ZS8	O81ZS8 plasmodium
15	101	14.0	651	10 09AW16	O9AW16 guillardia
16	99	13.7	1284	5 081AT3	O81AT3 plasmodium

17	98.5	13.6	371	17 058845	O58845 methanococc
18	98.5	13.6	1095	5 081K00	Q81K00 plasmodium
19	98	13.6	807	5 081625	Q81625 plasmodium
20	97.5	13.5	453	16 08RGF3	Q8RGF3 fusobacteri
21	97.5	13.5	841	2 09AIP1	Q9AIP1 carsonella
22	97.5	13.5	1811	5 081L76	Q81L76 plasmodium
23	96.5	13.3	1224	5 096209	O96209 plasmodium
24	96.5	13.3	1334	5 081KL3	Q81KL3 plasmodium
25	96.5	13.3	1979	5 081CX4	Q81CX4 plasmodium
26	96	13.3	529	16 08XW43	Q8XW43 clostridium
27	96	13.3	1571	5 081JF5	Q81JF5 plasmodium
28	95.5	13.2	1176	5 081B63	Q81B63 plasmodium
29	95.5	13.2	1451	5 081M56	Q81M56 plasmodium
30	95	13.1	265	11 054830	O54830 mus musculu
31	95	13.1	266	11 035334	O35334 mus musculu
32	95	13.1	435	16 08D345	Q8D345 wiggleswort
33	95	13.1	679	16 08EWF5	Q8EWF5 mycoplasma
34	95	13.1	1083	16 08D2T9	Q8D2T9 wiggleswort
35	95	13.1	1401	5 081S04	Q81S04 plasmodium
36	95	13.1	1807	5 081389	Q81389 plasmodium
37	95	13.1	2511	5 081L44	Q81L44 plasmodium
38	95	13.1	2651	5 081AV6	Q81AV6 plasmodium
39	95	13.1	4638	5 081K96	Q81K96 plasmodium
40	94.5	13.1	1000	5 081BX8	Q81BX8 plasmodium
41	94.5	13.1	1037	17 096ZS0	O96ZS0 sulfolobus
42	94.5	13.1	1132	5 081SC3	Q81SC3 plasmodium
43	94.5	13.1	1434	5 081CW7	Q81CW7 plasmodium
44	94.5	13.1	2510	5 094658	O94658 plasmodium
45	94.5	13.1	2695	5 081BR9	O81BR9 plasmodium

ALIGNMENTS

RESULT 1

O52975	PRELIMINARY;	PRT;	144 AA.
ID 052975			
AC 052975;			
DT 01-JUN-1998 (TREMBlrel. 06, Created)			
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)			
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)			
DE ORF-XI.			
OS Clostridium botulinum.			
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;			
OC Clostridium.			
OX NCBI_taxID=1491;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN-type E Iwanai;			
RX MEDLINE=98126542; PubMed=9465394;			
RA Kubota T., Yonekura N., Hariya Y., Isogai H., Amano K.,			
RA Fujii N.;			
RT "Gene arrangement in the upstream region of Clostridium botulinum type			
RT E and Clostridium butyricum Bl6340 progenitor toxin genes is different			
RT from that of other types.";			
RL FEMS Microbiol. Lett. 158:215-221(1998).			
DR EMBL; D88418; BAA24880.1;			
SQ SEQUENCE 144 AA; 16767 MW; 3A4F7DB7F67670BC CRC64;			

Query Match	100.0%;	Score 723;	DB 2;	Length 144;
Best Local Similarity	100.0%;	Pred. No. 5.7e-45;		
Matches 144;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MELQAFVFEDENLSSSGSIHLEKVKQKSPNYDYFKITFDYGLYTKNKGVLQKY	60	
Db	1	MELQAFVFEDENLSSSGSIHLEKVKQKSPNYDYFKITFDYGLYTKNKGVLQKY	60	
Qy	61	DLKNVISLVALKRDYLSLSLNKKQIKKFKNKKHKKFNLYVINEDIEKRTKNGIL	120	
Db	61	DLKNVISLVALKRDYLSLSLNKKQIKKFKNKKHKKFNLYVINEDIEKRTKNGIL	120	
Qy	121	EEVILNKMLLSILLGNEENLLOIS	144	

Fri Jul 25 09:41:20 2003

Db 121 EVILNKMILLSILLGNEENLLOIS 144

RESULT 2

O50600 PRELIMINARY: PRT: 142 AA.

AC O50600: (1)

DT 01-JUN-1998 (TREMBlrel. 06, Created)

DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)

DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE ORF-X1.

OS Clostridium botulinum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI_TaxID=1491;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-type F Langeland;

RA Kubota T., Yonekura N., Hariya Y., Isogai E., Isogai H., Amano K., Fujii N.;

RT "Gene arrangement in the upstream region of Clostridium botulinum type E and Clostridium butyricum BL6340 progenitor toxin genes is different from that of other types.";

RT FEMS Microbiol. Lett. 158:215-221(1998).

DR EMBL: AB004779; BAA24890.1;

SQ SEQUENCE 142 AA; 16396 MW; 8C0C682EB824122F CRC64;

Query Match 76.58; Score 553; DB 2; Length 142;

Best Local Similarity 73.94; Pred. No. 9.5e-33;

Matches 105; Conservative 19; Mismatches 18; Indels 0; Gaps 0;

Oy 3 LKQAFVFFDENLSSSSGSIHLEKVKQKSPNYDFKTFIDGYLYIKNKGVLKDYDL 62

Db 1 MKQTFSPFDDTLSSGSLIHLEKVRNRCSPNYQYFKIFIEGYLHIKNSGDIKLYDL 60

Oy 63 KNVISLVALKRDYLSLSNNKQIKKFNKKNHKNKFNLYVINEDIEKRIITKNGILEE 122

Db 61 KDLISLIALKADYLLKSSPNKKNPEFTNKNKHNLENRNFNLYVINEDINGKITKNGFLEE 120

Oy 123 VILNKMILLSILLGNEENLLOIS 144

Db 121 IILNRLLSILLGNEENLLOIA 142

Query Match 74.78; Score 540; DB 2; Length 142;

Best Local Similarity 72.59; Pred. No. 8.2e-32;

Matches 103; Conservative 20; Mismatches 19; Indels 0; Gaps 0;

RESULT 3

O50597 PRELIMINARY: PRT: 142 AA.

AC O50597: (1)

DT 01-JUN-1998 (TREMBlrel. 06, Created)

DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)

DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE ORF-X1.

OS Clostridium botulinum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI_TaxID=1491;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-type A Kyoto-F;

RA Kubota T., Yonekura N., Hariya Y., Isogai E., Isogai H., Amano K., Fujii N.;

RT "Gene arrangement in the upstream region of Clostridium botulinum type E and Clostridium butyricum BL6340 progenitor toxin genes is different from that of other types.";

RT FEMS Microbiol. Lett. 158:215-221(1998).

DR EMBL: AB004778; BAA24887.1;

SQ SEQUENCE 142 AA; 16396 MW; 0F8D08545A723EDB CRC64;

Query Match 74.78; Score 540; DB 2; Length 142;

Best Local Similarity 72.59; Pred. No. 8.2e-32;

Matches 103; Conservative 20; Mismatches 19; Indels 0; Gaps 0;

Oy 3 LKQAFVFFDENLSSSSGSIHLEKVKQKSPNYDFKTFIDGYLYIKNKGVLKDYDL 62

Db 1 MNQTFSPFDDTLSSGSLIHLEKVRNRCSPNYQYFKIFIEGYLHIKNSGDIKLYDL 60

Oy 63 KNVISLVALKRDYLSLSNNKQIKKFNKKNHKNKFNLYVINEDIEKRIITKNGILEE 122

Db 61 KDLISLIALKADYLLKSSPNKKNPEFTNKNKHNLENRNFNLYVINEDINGKITKNGILEE 120

Oy 123 VILNKMILLSILLGNEENLLOIS 144

Db 121 IILNRLLSILLGNEENLLOIA 142

RESULT 4

O81JP6 PRELIMINARY: PRT: 2279 AA.

ID O81JP6:

AC O81JP6: (1)

DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Hypothetical protein.

GN PF10_0146.

OS Plasmodium falciparum (isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=36329;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=3D7;

RA MEDLINE=22255705; PubMed=12368864;

RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W., K., Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.F., James K., Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S., Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S., Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B., Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A., McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C., Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W., Fraser C.M., Barrell B.;

RT "Genome sequence of the human malaria parasite Plasmodium falciparum.";

RL Nature 419:498-511(2002).

DR EMBL: AE014831; AAN35344.1;

KW Hypothetical protein.

SQ SEQUENCE 2279 AA; 267984 MW; E7D2B569E146C628 CRC64;

Query Match 15.4%; Score 111.5; DB 5; Length 2279;

Best Local Similarity 27.6%; Pred. No. 7.5;

Matches 34; Conservative 21; Mismatches 37; Indels 31; Gaps 4;

Oy 12 DENLSSSSGSIHLEKVKQKSPNYDFKTFIDGYLYIKNKGVLKDYDLK 63

Db 1040 DENSKEKSKSNEIKISEN-----IYIENKKKIELINTINNKNKKNK 1083

Oy 64 NVISL-----VALKRDYLSLSNNKQIKKFK---NIKKNHKNKFNLYVINEDIEKRIITK 116

Db 1084 RVIEIKKQAVPLKRNKNSVNSMDIKKVAQFKNNNNNNNHFKMSIPNYIFKKDIEIEIER 1143

Oy 117 NGI 119

Db 1144 KNI 1146

RESULT 5

O812L1 PRELIMINARY: PRT: 1782 AA.

ID O812L1:

AC O812L1: (1)

DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Hypothetical protein.

GN PF11470C.

OS Plasmodium falciparum (isolate 3D7).

[illegible]

```

QY 4 KOAFV-FEPDEN-----LSSSSGSIHLEKVKONSPNDYFKITFDGVLKRNKSGVI 56
Db 63 KENFVSFKDDKNLISVSPSTLKNLIFQDKNNLKK-----KVTIKGDIYRKKNDE-IY 116
QY 57 LKDYDLKNVISLVALKR-----DYLSSL-----SNKKOIKK-----FKNIKXHLKNKFNLYVI 106
Db 117 LDKNNLFNEINLLKKLINDVIRKDLFHKDNSIQIKPYIFDDYKKNVNIENKISIEII 176
QY 107 N-----EDIEKRITKNGILEV--ILNKMILLSILLGNEENLLQIS 144
Db 177 NFKKCNLENKKYNIHSSVVELRGTFNDELINLPILNETILOKIS 222

RESULT 9
QBI5T5
ID O815T5 PRELIMINARY; PRT: 2961 AA.
AC O815T5;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Hypothetical protein.
GN PFL0575w.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shalloom S.J., Suh B., Peterson J., Angiuoli S.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Vaidya A.B.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
DR EMBL; AE014846; AAN36204.1;
KW Hypothetical protein.
SQ SEQUENCE 2961 AA; 350583 MW; 1A887A5B8B5D0F69 CRC64;

Query Match 14.6%; Score 105.5; DB 5; Length 2961;
Best Local Similarity 25.0%; Pred. No. 26;
Matches 44; Conservative 29; Mismatches 54; Indels 49; Gaps 8;

QY 7 FVEFDENLSSSSGSIHLEKVKON--CSPNDY----YFKITFDGY-----LYIK 50
Db 342 YVHDLSENKYSSLPPEKKYKENIKCMKNFEKSHNMNKIOYVEDFOKENTENGKILYLK 401
QY 51 NKSGVILDKYDLK-----NVISLVALKRDYLSLSLNKKOIKKFKNIKXHLK 98
Db 402 ERKGI--NKYNEKDMCMCHSAFDNMVQ-----NNVLDNMCNKKNIKK--KMKKNVNIK 454
QY 99 NKFNLYVINEDEIKRITKNGILEE-----VILNKMILLSILLGNEENL 140
Db 455 FWDVNYVINDVSNKMHNNHNSLOENNYSVLOIPNNKQINNIRANVILNNEFN 510

RESULT 10
O25314
ID O25314 PRELIMINARY; PRT: 1001 AA.
AC O25314;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE Hypothetical protein.
GN HP0592.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

```

```

OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
DR EMBL; AE000572; AAD07657.1;
DR TIGR; HP0592;
DR InterPro; IPR006935; ResIII.
DR Pfam; PF04851; ResIII; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1001 AA; 114809 MW; 1817F729A96F48B1 CRC64;

Query Match 14.5%; Score 104.5; DB 16; Length 1001;
Best Local Similarity 37.0%; Pred. No. 11;
Matches 34; Conservative 12; Mismatches 31; Indels 15; Gaps 4;

QY 24 LEKVONCSPNDYFKITF---IDGYL--IKNKGVLKDYDLKNVISLVALKRDYLSL 78
Db 118 LESVQONILRLTDYFLEFKRHUKTYTEGVKSPSNII--NHYIKN-----ODELSV 167
QY 79 SLSNKKQIKKFKNIKXHLKNKFNLYVINEDI 110
Db 168 LLTNSATDKEGNILKNSENLFNTKSIFENI 199

RESULT 11
QBI177
ID QBI177 PRELIMINARY; PRT: 1226 AA.
AC QBI177;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Hypothetical protein.
GN PFI4_0371.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shalloom S.J., Suh B., Peterson J., Angiuoli S.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Vaidya A.B.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
DR EMBL; AE014822; AAN36984.1;
KW Hypothetical protein.
SQ SEQUENCE 1226 AA; 148166 MW; E3F53927D1F73E64 CRC64;

Query Match 14.5%; Score 104.5; DB 5; Length 1226;
Best Local Similarity 22.2%; Pred. No. 13;
Matches 42; Conservative 39; Mismatches 53; Indels 55; Gaps 8;

```

```

QY 4 KOAFVFEF-----DENLSSSGSIHLEKVKONCSNPYDFKTF--IDGYLYI-- 49
DB 724 KKKEVREICSEQNKIINDSIKFFSNNEYKNIKENC--DLLKLTITFOIKTMIVFN 781
QY 50 -----KNKSGVLDKYNISLVALKRDYLSLSLNKKQIKFKNKNKHL----- 97
DB 782 LIQIKYKND--KEKYLRLNIIISILLDLSSYSISNLNHTSYMLNENLIISHFLYFCF 838
QY 98 ----KNKFNLYVIN-----EDIEKRYT---KNGILEEVLNKMLLSL 133
DB 839 SELYNPKYNEFIIDYKKTHTSOTCNFNFSIENIKSEITKKKKSIKQMLNEDILLNF 898
QY 134 LGNEENLLQ 142
DB 899 YADYKNYMQ 907

RESULT 12
QBTJRO
ID QBTJRO PRELIMINARY; PRT; 245 AA.
AC QBTJRO
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE Predicted protein.
GN MA3719.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity."
RL Genome Res. 12:532-542(2002).
DR EMBL; AE011083; AAM07072.1;
DR InterPro; IPR001638; SBP_bac_3.
DR PROSITE; PS01039; SBP_BACTERIAL_3.1.
KW Complete proteome.
SQ SEQUENCE 245 AA; 28108 MW; A8B44EEC7105276C CRC64;

Query Match 14.4%; Score 104; DB 17; Length 245;
Best Local Similarity 28.4%; Pred. No. 3.1;
Matches 40; Conservative 22; Mismatches 43; Indels 36; Gaps 8;

QY 30 NCSNPYDFKITFDGYLYTKNKGVLTKYDLKKNVISLVALK-----RDYLSL 78
DB 8 NCS-NYSKYAVS-----CMYTN--SILSSVPLKNLPNLAQNDVLFLIDLGDYTL 60
QY 79 SLSNKKQIKFKNKNKLNKENLYVINEDIEKRI-----YKNGILE-----EV- 123
DB 61 ECLNRSERHFELKLTAYPKKR---YVSVTVLKHILCOLLNKNSVLQIATYKDFGEVH 117
QY 124 ILNKMLLSLLGNEENLLQIS 144
DB 118 ILNHNELHICLSYSENATLS 138

RESULT 13

```

```

Q8ICK6
ID Q8ICK6 PRELIMINARY; PRT; 227 AA.
AC Q8ICK6
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Hypothetical protein.
GN MAL6P1.263.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Cherevach I., Davis P., Goodhead I., Stevens K., Muncall K.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.;
RL Submitted (SFP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844505; CAD50435.1;
KW Hypothetical protein.
SQ SEQUENCE 227 AA; 26900 MW; FAC5DB456CD9BB28 CRC64;

Query Match 14.2%; Score 102.5; DB 5; Length 227;
Best Local Similarity 27.0%; Pred. No. 3.7;
Matches 33; Conservative 26; Mismatches 32; Indels 31; Gaps 4;

QY 42 FIDGYLY---IKNKGVLTKYDLKKNVISLVALKRDYLSLSLNKKQIKFKNKNKHLK 98
DB 93 FLNGLYLFSRIGHNINLLNE-----LSSIKKDFELNKKTRKIKKKFQVKNKNI 144
QY 99 NKENLYVINE-----DIEKRTKNGILEEVLNKMLLSLILGNEENL 140
DB 145 YKLLQDIKEKKRKRKEIYEKLSIEIKKLDVKVQEKIEKDEINNLHKL--NIENI 202
QY 141 LQ 142
DB 203 IK 204

RESULT 14
Q812S8
ID Q812S8 PRELIMINARY; PRT; 4202 AA.
AC Q812S8
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Hypothetical protein.
GN PF1120C.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2225708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Muncall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Felwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes S., Larkin N., Lawson D., Lemard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Sulston J.E., Craig A., Unwin L., Whitehead S., Woodward J.,
RA Taylor K., Thayer A., Newbold C., Barrell B.G.
RT "Sequence of plasmodium falciparum chromosomes 1, 3-9 and 13."
RL Nature 419:527-531(2002).
DR EMBL; AL929357; CAD51910.1;
KW Hypothetical protein.
SQ SEQUENCE 4202 AA; 509737 MW; 16EEE070C3F1DA0E CRC64;

```

Search completed: July 24, 2003, 20:08:41
Job time : 112.403 secs

Query Match 14.1%; Score 102; DB 5; Length 4202;
Best Local Similarity 27.1%; Pred. No. 65;
Matches 39; Conservative 26; Mismatches 51; Indels 28; Gaps 7;
QY 12 DENLSSSGS-----THLEKVKONCSPNYDYFKITFDIGLYIKNKSGVI---LDKYDLK 63
DB 3001 DENTQGRDESINFLLIHLKSGSKYFNSVTEYIKIVNQKKQYINNDTSILKKKIPQHDFY 3060
QY 64 NVISVALKRDYLSLSL-----NNQIKKFKNKKHL-KNKFNLVYVINEDIE---KRI 114
DB 3061 N-----ICYKNEYITNFMYNINLERKKYIRINKKHIDKKKYYYYRGRDKIGQORKRL 3116
QY 115 TKMGILEEVILNKMLLSILLGNEE 138
DB 3117 -----KLYFSKPASSYILKNKE 3133

RESULT 15

Q9AW16 PRELIMINARY; PRT; 651 AA.
AC Q9AW16:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE DNA primase.
GN DNAG.
OS Guillardia theta (Cryptomonas phi).
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
OX NCBI_TaxID=55529;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20087226; PubMed=10618395;
RA Zauner S., Fraunholz M., Wastl J., Penny S.L., Beaton M.,
Cavallier-Smith T., Maier U., Douglas S.;
RT "Chloroplast protein and centrosomal genes, a tRNA intron, and odd
telomeres in an unusually compact eukaryotic genome, the cryptomonad
nucleomorph.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:200-205(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2123349; PubMed=11323671;
RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.,
Wu X., Reith M., Cavallier-Smith T., Maier U.;
RT "The highly reduced genome of an enslaved algal nucleus.";
RL Nature 410:1091-1096(2001).
DR EMBL; AJ010592; CAC27055.1; -;
DR HSSP; Q9X4D0; ID00.
DR InterPro; IPR006647; Toprim_primase.
DR InterPro; IPR006154; Toprim_sub.
DR InterPro; IPR002694; Znf_CHC2.
DR Pfam; PF01807; zf-CHC2; 1.
DR ProDom; PD002276; Toprim_primase; 1.
DR ProDom; PD002988; Znf_CHC2; 1.
DR SMART; SM00493; TOPRIN; 1.
SQ SEQUENCE 651 AA; 77961 MW; 00F0EB569058F3B CRC64;

Query Match 14.0%; Score 101; DB 10; Length 651;
Best Local Similarity 28.9%; Pred. No. 13;
Matches 50; Conservative 27; Mismatches 56; Indels 40; Gaps 11;
QY 2 ELKQAFVEFDENLSSSGSI-----HLEKVKO--NCSPNYDYFKITFDIGLYIKNKSG 54
DB 167 ETFYMLYKFP-RNLSNLTWKILNLGYSNKKISOIINCQENHEYRL-----LKTMTKY 219
QY 55 VILDKYDLKNVI-----SLVALKRDYLSLSL-----NNQIKKFKNIK-NKHLKNKF 101
DB 220 VFEDKINKKVVIRDFINERLVIPIRNYLCITGLGARTINSKKIPKYLNSKENKFFKKK 279
QY 102 NLYVINEDIEKRITK-----NGILEEVL--NKMLLSIL-LGNEENLQI 143
DB 280 ILF--SEEIKSNLSLKKCLICEGYLDSTITLFGNGIRFVSASLGSSNNFQI 330



results of **BLAST**

BLASTP 2.2.6 [Apr-09-2003]

Reference:

Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

RID: 1059158413-09039-19779

Query=

(204 letters)

Database: All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF

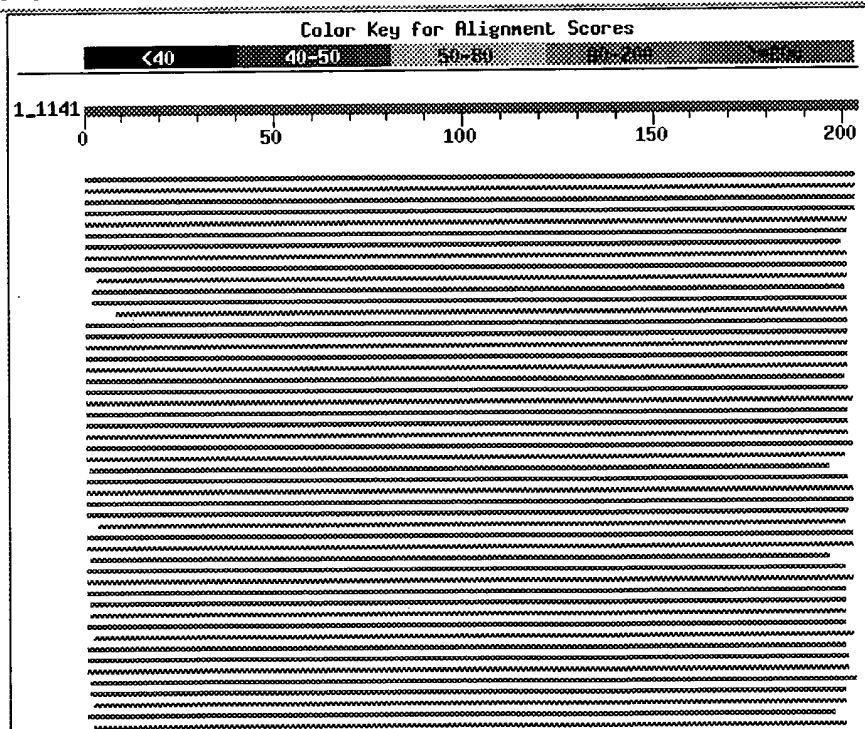
1,475,854 sequences; 475,827,238 total letters

If you have any problems or questions with the results of this search please refer to the [BLAST FAQs](#)

[Taxonomy reports](#)

Distribution of 100 Blast Hits on the Query Sequence

Mouse-over to show defline and scores. Click to show alignments



[Related Structures](#)

Score E

Sequences producing significant alignments:

(bits) Value

gi	134649	sp	P28760	SODM BACCA	Superoxide dismutase [Mn] >g...	373	e-103
gi	134650	sp	P00449	SODM BACST	Superoxide dismutase [Mn] >g...	372	e-102
gi	7798571	dbj	EAA95631.1		superoxide dismutase [Geobacillu...	369	e-101
gi	448837	prf	1918164A		superoxide dismutase	367	e-101
gi	15613972	ref	NP_242275.1		superoxide dismutase [Bacillus...	311	4e-84
gi	16079557	ref	NP_390381.1		superoxide dismutase [Bacillus...	308	3e-83
gi	3355882	dbj	EAA31974.1		superoxide dismutase [Bacillus s...	308	3e-83
gi	32468794	emb	CAB14432.2		superoxide dismutase [Bacillus ...	307	7e-83
gi	3790114	emb	CAA05291.1		manganese superoxide dismutase [...	306	1e-82
gi	30264347	ref	NP_846724.1		superoxide dismutase, Mn [Baci...	306	2e-82
gi	23200182	pdb	1JR9	A	Chain A, Crystal Structure Of Mangan...	303	1e-81
gi	23099387	ref	NP_692853.1		manganese superoxide dismutase...	302	3e-81
gi	30022355	ref	NP_833986.1		Superoxide dismutase [Mn] [Bac...	298	2e-80
gi	16800546	ref	NP_470814.1		superoxide dismutase [Listeria...	285	4e-76
gi	29375094	ref	NP_814247.1		superoxide dismutase, Mn [Ente...	282	2e-75
gi	16803479	ref	NP_464964.1		superoxide dismutase [Listeria...	282	2e-75
gi	7433329	pir	S20019		superoxide dismutase (EC 1.15.1.1) (...	281	5e-75
gi	134666	sp	P28763	SODM LISIV	Superoxide dismutase [Mn] >g...	276	1e-73
gi	3320377	gb	AAC26483.1		manganese superoxide dismutase [V...	276	2e-73
gi	18378721	gb	AAL68691.1	AF462457.1	superoxide dismutase [...	268	5e-71
gi	32492801	gb	AAP85516.1		SodA [Aeromonas salmonicida subs...	264	6e-70
gi	8977980	emb	CAB95744.1		superoxide dismutase [Staphyloco...	264	7e-70
gi	27468158	ref	NP_764795.1		superoxide dismutase SodA [Sta...	263	1e-69
gi	15924543	ref	NP_372077.1		superoxide dismutase [Staphylo...	263	2e-69
gi	15599664	ref	NP_253158.1		superoxide dismutase [Pseudomo...	262	3e-69
gi	11121341	emb	CAC14833.1		superoxide dismutase [Staphyloc...	261	3e-69
gi	17227566	ref	NP_484114.1		superoxide dismutase [Nostoc s...	260	8e-69
gi	23135202	ref	ZP_00116957.1		hypothetical protein [Cytoph...	258	4e-68
gi	1084032	pir	A53294		superoxide dismutase (EC 1.15.1.1) (...	257	1e-67
gi	21231716	ref	NP_637633.1		superoxide dismutase [Xanthomo...	255	3e-67
gi	31074373	gb	AAP41921.1		Mn/Fe superoxide dismutase [Chlo...	255	3e-67
gi	23126478	ref	ZP_00108372.1		hypothetical protein [Nostoc...	255	3e-67
gi	21243120	ref	NP_642702.1		superoxidase dismutase [Xantho...	254	4e-67
gi	1150790	gb	AAB47971.1		superoxide dismutase [Xanthomonas...	254	5e-67
gi	23200074	pdb	1GV3	A	Chain A, The 2.0 Angstrom Resolution...	254	6e-67
gi	1711423	sp	P50059	SOD2 PLEBO	Superoxide dismutase [Mn] 2...	251	6e-66
gi	1711456	sp	P09214	SODM THETH	Superoxide dismutase [Mn] >...	251	6e-66
gi	23061777	ref	ZP_00086602.1		hypothetical protein [Pseudo...	251	7e-66
gi	77595	pir	S07147		superoxide dismutase (EC 1.15.1.1) (Mn...	251	7e-66
gi	23470984	ref	ZP_00126316.1		hypothetical protein [Pseudo...	251	7e-66
gi	26987682	ref	NP_743107.1		superoxide dismutase (Mn) [Pse...	250	1e-65
gi	134640	sp	P23744	SODF METJ	Superoxide dismutase [Mn-Fe] ...	249	2e-65
gi	2500831	sp	P77929	SODM PSEPU	Superoxide dismutase [Mn] >...	248	4e-65
gi	32474407	ref	NP_867401.1		superoxide dismutase, Mn famil...	246	1e-64
gi	15806297	ref	NP_295003.1		superoxide dismutase (sodA), M...	246	2e-64
gi	1711422	sp	P50058	SOD1 PLEBO	Superoxide dismutase [Mn] 1...	244	5e-64
gi	1711455	sp	P53653	SODM THEAQ	Superoxide dismutase [Mn] >...	244	6e-64
gi	30023475	ref	NP_835106.1		Superoxide dismutase [Mn] [Bac...	243	1e-63
gi	21397929	ref	NP_653914.1		sodfe_C, Iron/manganese supero...	242	2e-63
gi	22996085	ref	ZP_00040357.1		hypothetical protein [Xylell...	242	3e-63
gi	22995080	ref	ZP_00039563.1		hypothetical protein [Xylell...	241	6e-63
gi	24182472	gb	AAN16456.2		manganese-containing superoxide ...	240	8e-63
gi	32035144	ref	ZP_00135190.1		hypothetical protein [Actino...	239	2e-62
gi	28199854	ref	NP_780168.1		superoxide dismutase [Xylella ...	238	4e-62
gi	15923123	ref	NP_370657.1		superoxide dismutase [Staphylo...	237	8e-62
gi	16329169	gb	AAG44813.2	AF273269.1	superoxide dismutase [...	237	9e-62
gi	23040106	ref	ZP_00071638.1		hypothetical protein [Tricho...	236	1e-61
gi	1711438	sp	P53642	SODM BORPE	Superoxide dismutase [Mn] >...	236	2e-61
gi	15839203	ref	NP_299891.1		superoxide dismutase [MN] [Xyl...	235	4e-61
gi	16124175	ref	NP_407488.1		superoxide dismutase [Mn] [Yer...	234	4e-61
gi	1711458	sp	P53655	SODM YEREN	Superoxide dismutase [Mn] >...	234	7e-61
gi	21910607	ref	NP_664875.1		superoxide dismutase [Mn] [Str...	234	9e-61
gi	541376	pir	PN0614		superoxide dismutase (EC1.15.1.1) (M...	233	1e-60

gi	23130308	ref	ZP_00112125.1	hypothetical protein [Nostoc...	233	1e-60	
gi	6094321	sp	O30826	SODM HAEDU Superoxide dismutase [Mn] >...	233	2e-60	
gi	33151556	ref	NP_872909.1	manganese superoxide dismutase...	233	2e-60	
gi	19746368	ref	NP_607504.1	superoxide dismutase [Streptoc...	232	2e-60	
gi	13959576	sp	P77957	SODM STRPY Superoxide dismutase [Mn]	232	3e-60	
gi	15804498	ref	NP_290538.1	superoxide dismutase, manganes...	232	3e-60	
gi	24115202	ref	NP_709712.1	superoxide dismutase, manganes...	232	3e-60	
gi	15551671	emb	CAC69393.1	manganese superoxide dismutase...	231	5e-60	
gi	15675326	ref	NP_269500.1	superoxide dismutase (Fe/Mn) [...]	231	5e-60	
gi	16131748	ref	NP_418344.1	superoxide dismutase, manganes...	231	5e-60	
gi	2808462	emb	CAA11227.1	superoxide dismutase [Streptococ...	230	9e-60	
gi	26250673	ref	NP_756713.1	Superoxide dismutase [Mn] [Esc...	230	1e-59	
gi	28871594	ref	NP_794213.1	superoxide dismutase, Mn [Pseu...	230	1e-59	
gi	16762379	ref	NP_457996.1	manganese superoxide dismutase...	230	1e-59	
gi	16767321	ref	NP_462936.1	superoxide dismutase, manganes...	229	2e-59	
gi	15672390	ref	NP_266564.1	superoxide dismutase [Lactococ...	229	2e-59	
gi	3212626	pdb	1VEW	A Chain A, Manganese Superoxide Dismuta...	229	2e-59	
gi	808038	gb	AAC43331.1	Mn-superoxide dismutase	228	3e-59	
gi	13399622	pdb	1I0H	A Chain A, Crystal Structure Of The E....	228	5e-59	
gi	14719524	pdb	1EN5	A Chain A, Crystal Structure Analysis ...	228	5e-59	
gi	2500830	sp	Q59679	SODM PASHA Superoxide dismutase [Mn] >...	227	7e-59	
gi	32029513	ref	ZP_00132526.1	hypothetical protein [Haemop...	227	7e-59	
gi	541031	pir	PN0615	superoxide dismutase (EC 1.15.1.1) (F...	226	1e-58	
gi	14719520	pdb	1EN4	A Chain A, Crystal Structure Analysis ...	226	1e-58	
gi	14719528	pdb	1EN6	A Chain A, Crystal Structure Analysis ...	226	1e-58	
gi	22135437	gb	AAM93199.1	superoxide dismutase [Thermus an...	226	1e-58	
gi	22135443	gb	AAM93202.1	superoxide dismutase [Thermus fi...	226	2e-58	
gi	13399618	pdb	1I08	A Chain A, Crystal Structure Analysis ...	226	2e-58	
gi	17226292	gb	AAL37717.1	AF413524.1 superoxide dismutase [...]	226	2e-58	
gi	22135445	gb	AAM93203.1	superoxide dismutase [Thermus os...	224	4e-58	
gi	3334336	sp	Q42684	SODM CHLRE Superoxide dismutase [Mn], ...	224	6e-58	
gi	5758312	gb	AAD50778.1	AF162664.1 manganese co-factored s...	224	6e-58	
gi	22135439	gb	AAM93200.1	superoxide dismutase [Thermus br...	224	7e-58	
gi	22135441	gb	AAM93201.1	superoxide dismutase [Thermus ig...	224	8e-58	
gi	22135447	gb	AAM93204.1	superoxide dismutase [Thermus sc...	224	9e-58	
gi	15900660	ref	NP_345264.1	superoxide dismutase, manganes...	224	9e-58	
gi	22536952	ref	NP_687803.1	superoxide dismutase, Fe-Mn [S...	223	1e-57	

Alignments

Get selected sequences

Select all

Deselect all

☒ >gi|134649|sp|P28760|SODM_BACCA Superoxide dismutase [Mn]
gi|98098|pir|S22053 superoxide dismutase (EC 1.15.1.1) (Mn) - Bacillus caldotena
gi|39453|emb|CAA44556.1 Manganese superoxide dismutase [Bacillus caldotenax]
gi|7578509|gb|AAF64074.1|AF147780.1 superoxide dismutase [Geobacillus thermoleovo
gi|384175|prf|1905285A superoxide dismutase

Length = 204

Score = 373 bits (958), Expect = e-103

Identities = 189/204 (92%), Positives = 189/204 (92%)

Query: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGH PDLQNKXXXXXXXXX 60
MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGH PDLQNK
Sbjct: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGH PDLQNK SLEELLSN 60

Query: 61 XXXXXXIRTAVRNNGGGHANHSFWTILSPNGGGPEPTGELAEAINKKFGSFTAFKDEFS 120
IRTAVRNNGGGHANHSFWTILSPNGGGPEPTGELAEAINKKFGSFTAFKDEFS
Sbjct: 61 LEALPESIRTAVRNNGGGHANHSFWTILSPNGGGPEPTGELAEAINKKFGSFTAFKDEFS 120

Query: 121 KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE 180
KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE

Sbjct: 121 KAAAGRFGSGAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE 180

Query: 181 YIAAFWNIVNWDEVAKRYSEAKAK 204

YIAAFWNIVNWDEVAKRYSEAKAK

Sbjct: 181 YIAAFWNIVNWDEVAKRYSEAKAK 204

☐ >gi|134650|sp|P00449|SODM_BACST Superoxide dismutase [Mn]
 gi|1070456|pir|DSBSNF superoxide dismutase (EC 1.15.1.1) (Mn) - Bacillus
 stearothermophilus
 gi|143203|gb|AAA22600.1 manganese superoxide dismutase (EC 1.15.1.1)
 gi|143552|gb|AAA22765.1 Mn-superoxide dismutase
 gi|143555|gb|AAA22767.1 Mn-superoxide dismutase
 Length = 204

Score = 372 bits (954), Expect = e-102

Identities = 187/204 (91%), Positives = 189/204 (92%)

Query: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKXXXXXXXXX 60
 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNK

Sbjct: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKSLLELLSN 60

Query: 61 XXXXXXI RTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFS 120
 IRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELA+AINKKFGSFTAFKDEFS

Sbjct: 61 LEALPESIRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELADAINKKFGSFTAFKDEFS 120

Query: 121 KAAAGRFGSGAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE 180
 KAAAGRFGSGAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE

Sbjct: 121 KAAAGRFGSGAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE 180

Query: 181 YIAAFWNIVNWDEVAKRYSEAKAK 204

YIAAFWN+VNWDEVAKRYSEAKAK

Sbjct: 181 YIAAFWNVNWDEVAKRYSEAKAK 204

☐ >gi|7798571|dbj|BAA95631.1 superoxide dismutase [Geobacillus thermoleovorans]
 Length = 204

Score = 369 bits (946), Expect = e-101

Identities = 187/204 (91%), Positives = 187/204 (91%)

Query: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKXXXXXXXXX 60
 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNK

Sbjct: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKSLLELLSN 60

Query: 61 XXXXXXI RTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFS 120
 IRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFS

Sbjct: 61 LEALPESIRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFS 120

Query: 121 KAAAGRFGSGAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE 180
 KAAAGRF GAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE

Sbjct: 121 KAAAGRFSPGAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE 180

Query: 181 YIAAFWNIVNWDEVAKRYSEAKAK 204

YIAAFWNIVNWDEVAKRYSEAKAK

Sbjct: 181 YIAAFWNIVNWDEVAKRYSEAKAK 204

☐ >gi|448837|prf||1918164A superoxide dismutase
 Length = 205

Score = 367 bits (941), Expect = e-101

Identities = 187/205 (91%), Positives = 189/205 (92%), Gaps = 1/205 (0%)

Query: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKXXXXXXXXX 60
 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNK

Sbjct: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKSLLELLSN 60

Query: 61 XXXXXXIIRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFS 120
 IRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELA+AINKKFGSFTAFKDEFS
 Sbjct: 61 LEALPESIRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELADAINKKFGSFTAFKDEFS 120

Query: 121 KAAAGRFGSGWAWLVVNNGELEITSTPNQDS-PIMEGKTPILGLDVWEHAYYLKYQNRNP 179
 KAAAGRFGSGWAWLVVNNGELEITSTPNQDS PIMEGKTPILGLDVWEHAYYLKYQNRNP
 Sbjct: 121 KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRNP 180

Query: 180 EYIAAFWNIVNWDEVAKRYSEAKAK 204
 EYIAAFWN+VNWDEVAKRYSEAKAK
 Sbjct: 181 EYIAAFWNVNWDEVAKRYSEAKAK 205

☐ >gi|15613972|ref|NP_242275.1| superoxide dismutase [Bacillus halodurans]
 gi|25285807|pir|A83826 superoxide dismutase sodA [imported] - Bacillus haloduran
 C-125)
 gi|10174025|dbj|BAB05128.1| superoxide dismutase [Bacillus halodurans]
 Length = 202

Score = 311 bits (798), Expect = 4e-84
 Identities = 153/202 (75%), Positives = 169/202 (83%)

Query: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKXXXXXXXXX 60
 M FELP LPYP +ALEPHID+ TMNIHH KHHNTYVT L NAALEGH L K
 Sbjct: 1 MAFELPKLPYPANALEPHIDEATMNIHHGKHNTYVTKL NAALEGHSALAEKSIEALVSD 60

Query: 61 XXXXXXIIRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFS 120
 IRTAVRNNGGGHANH+LFW ILSPNGGG PTGELA+AIN +FGSF FK++F+
 Sbjct: 61 LDAVPENIRTAVRNNGGGHANHTLFWQILSPNGGGAPT GELADAINAEFGSFDQFKEKFA 120

Query: 121 KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRNP 180
 AAA RFGSGWAWLVVN+G+LEITSTPNQD+P+MEGKTPILGLDVWEHAYYL YQNRNP+
 Sbjct: 121 DAAANRFGSGWAWLVVNDGKLEITSTPNQDTPLEMGKTPILGLDVWEHAYYLNYQNRNP 180

Query: 181 YIAAFWNIVNWDEVAKRYSEAK 202
 YI+AFWN+VNWDEVAKRY+EAK
 Sbjct: 181 YISAFWNVNWDEVAKRYNEAK 202

☐ >gi|16079557|ref|NP_390381.1| superoxide dismutase [Bacillus subtilis]
 gi|1711437|sp|P54375|SODM BACSU Superoxide dismutase [Mn] (General stress protein
 gi|7433327|pir|B69709 superoxide dismutase (EC 1.15.1.1) (Mn) sodA - Bacillus su
 gi|1303851|dbj|EAA12507.1| YqgD [Bacillus subtilis]
 Length = 226

Score = 308 bits (790), Expect = 3e-83
 Identities = 152/200 (76%), Positives = 166/200 (83%)

Query: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKXXXXXXXXX 60
 M +ELP LPY YDALEPHIDKETM IHHTKHHNTYVTNLN A+EG+ L NK
 Sbjct: 1 MAYELPELPYAYDALEPHIDKETMTIHHTKHHNTYVTNLNKAVEGNTALANKSVEELVAD 60

Query: 61 XXXXXXIIRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFS 120
 IRTAVRNNGGGHANH LFWT+LSPNGGGEPTG LAE IN FGSF FK++F+
 Sbjct: 61 LDSVPENIRTAVRNNGGGHANHKLFWTLLSPNGGGEPTGALAEINSVFGSFDKFKEQFA 120

Query: 121 KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRNP 180
 AAAGRFGSGWAWLVVNNG+LEITSTPNQDSP+ EGKTPILGLDVWEHAYYL YQNRNP+
 Sbjct: 121 AAAAGRFGSGWAWLVVNNGKLEITSTPNQDSPLESGKTPILGLDVWEHAYYLNYQNRNP 180

Query: 181 YIAAFWNIVNWDEVAKRYSE 200
 YI+AFWN+VNWDEVA+ YS+
 Sbjct: 181 YISAFWNVNWDEVARLYSD 200

☐ >gi|3355382|dbj|BAA31974.1| superoxide dismutase [Bacillus subtilis]
 Length = 202

Score = 308 bits (790), Expect = 3e-83
Identities = 155/202 (76%), Positives = 168/202 (83%)

Query: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNXXXXXXXXX 60
M +ELP LPY YDALEPHIDKETM IHHTKHHNTYVTNLN A+EG+ L NK
Sbjct: 1 MAYELPELPYAYDALEPHIDKETMTIHHTKHHNTYVTNLNKAVEGNTALANKSVEELVAD 60

Query: 61 XXXXXXXXIRTAVRNNGGGHANHSLFWTILSPNGGGGEPTGELAEAINKKFGSFTAFKDEFS 120
IRTAVRNNGGGHANH LFWT+LSPNGGGGEPTG LAE IN FGSF FK++F+
Sbjct: 61 LDSVPENIRTAVRNNGGGHANHKLFWTLLSPNGGGGEPTGALAEIEINSVFGSFDKFKEQFA 120

Query: 121 KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE 180
AAAGRFGSGWAWLVVNNG+LEITSTPNQDSP+ EGKTPILGLDVWEHAYYL YQNRPE+
Sbjct: 121 AAAAGRFGSGWAWLVVNNGKLEITSTPNQDSPLESGKTPILGLDVWEHAYYLYQNRPE 180

Query: 181 YIAAFWNIVNWDEVAKRYSEAK 202
YI+AFWN+VNWDEVA+ YSEAK
Sbjct: 181 YISAFWNVNWDEVARLYSEAK 202

☐ >gi|32468794|emb|CAB14432.2| superoxide dismutase [Bacillus subtilis subsp. subt
Length = 202

Score = 307 bits (786), Expect = 7e-83
Identities = 154/202 (76%), Positives = 167/202 (82%)

Query: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNXXXXXXXXX 60
M +ELP LPY YDALEPHIDKETM IHHTKHHNTYVTNLN A+EG+ L NK
Sbjct: 1 MAYELPELPYAYDALEPHIDKETMTIHHTKHHNTYVTNLNKAVEGNTALANKSVEELVAD 60

Query: 61 XXXXXXXXIRTAVRNNGGGHANHSLFWTILSPNGGGGEPTGELAEAINKKFGSFTAFKDEFS 120
IRTAVRNNGGGHANH LFWT+LSPNGGGGEPTG LAE IN FGSF FK++F+
Sbjct: 61 LDSVPENIRTAVRNNGGGHANHKLFWTLLSPNGGGGEPTGALAEIEINSVFGSFDKFKEQFA 120

Query: 121 KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE 180
AAAGRFGSGWAWLVVNNG+LEITSTPNQDSP+ EGKTPILGLDVWEHAYYL YQNRPE+
Sbjct: 121 AAAAGRFGSGWAWLVVNNGKLEITSTPNQDSPLESGKTPILGLDVWEHAYYLYQNRPE 180

Query: 181 YIAAFWNIVNWDEVAKRYSEAK 202
YI+AFWN+VNWDEVA+ YSE K
Sbjct: 181 YISAFWNVNWDEVARLYSERK 202

☐ >gi|3790114|emb|CAA05291.1| manganese superoxide dismutase [Bacillus licheniform
Length = 202

Score = 306 bits (785), Expect = 1e-82
Identities = 155/202 (76%), Positives = 167/202 (82%)

Query: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNXXXXXXXXX 60
M ++LP LPY YDALEPHIDKETMNIHHTKHHNTYVT LN A+ G DL++K
Sbjct: 1 MAYKLPELPYAYDALEPHIDKETMNIHHTKHHNTYVTNKLNEAVAGKQDLESKSVEELVAN 60

Query: 61 XXXXXXXXIRTAVRNNGGGHANHSLFWTILSPNGGGGEPTGELAEAINKKFGSFTAFKDEFS 120
IRTAVRNNGGGHANHSLFW +LSPNGGG PTGELAEAIN KFGSF FK++F+
Sbjct: 61 LDAVPENIRTAVRNNGGGHANHSLFWKLLSPNGGGAPTGELAEAINSKFGSFDQFKEDFA 120

Query: 121 KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE 180
AAA RFGSGWAWLVVNNGELEITSTPNQDSP+ EGKTPILGLDVWEHAYYL YQNRPE+
Sbjct: 121 AAAAARFGSGWAWLVVNNGELEITSTPNQDSPLESGKTPILGLDVWEHAYYLYQNRPE 180

Query: 181 YIAAFWNIVNWDEVAKRYSEAK 202
YI AFWN+VNWDEVA YSEAK
Sbjct: 181 YIKAFWNVNWDEVAPLYSEAK 202

☐ >gi|30264347|ref|NP_846724.1| superoxide dismutase, Mn [Bacillus anthracis str.
gi|30259005|gb|AAP28210.1| superoxide dismutase, Mn [Bacillus anthracis str. Ames

Length = 203

Score = 306 bits (783), Expect = 2e-82



Identities = 149/199 (74%), Positives = 164/199 (82%)

Query: 4 ELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKXXXXXXXXXXXX 63
 ELP LPY YDALEPH DKETMNIHHTKHHNTY+TNLNAALEGH +L +K
 Sbjct: 5 ELPNLPYAYDALEPHFDKETMNIHHTKHHNTYITNLNAALEGHAELADKSVEELVANLNE 64

Query: 64 XXXXIIRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFKAA 123
 IRTAVRNNGGGHANH+ FWITILSPNGGG+P GELA AI KFGSF AFK+EF+KA
 Sbjct: 65 VPEAIRTAVRNNGGGHANHTFFWTILSPNGGQPVGELATAIEAKFGSFDKKEEFAKAG 124

Query: 124 AGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRREPEYIA 183
 A RFGSGWAWLVVNNGELE+TSTPNQDSP+ EGKTP++GLDVWEHAYYL YQNRREPEYI
 Sbjct: 125 ATRFGSGWAWLVVNNGELEVTSTPNQDSPLTEGKTPVIGLDVWEHAYYLNQNRREPEYI 184

Query: 184 AFWNIVNWDEVAKRYSEAK 202
 AFWN+V+W+ KRY EAK
 Sbjct: 185 AFWNVVDWNAEKRYQEAK 203

 >gi|23200182|pdb|1JR9|A  Chain A, Crystal Structure Of Manganese Superoxide Dism
 Bacillus Halodenitrificans
 Length = 202

Score = 303 bits (776), Expect = 1e-81


Identities = 149/199 (74%), Positives = 168/199 (84%)

Query: 3 FELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKXXXXXXXXXXXX 62
 FELP LPY YDALEP IDKETMNIHHTKHHNTYVT LN ALEGH DL+NK
 Sbjct: 3 FELPELPYAYDALEPTIDKETMNIHHTKHHNTYVTKLNGALEGHEDLKNKSLNDLISNLD 62

Query: 63 XXXXIIRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFKAA 122
 IRTAVRNNGGGHANHSLFW ++SPNGGG+PTGE+A+ IN K+GSF F++EF+ A
 Sbjct: 63 AVPENIRTAVRNNGGGHANHSLFWKLMSPNGGKPTGEVADKINDKYGSFEKFQEEFAAA 122

Query: 123 AAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRREPEYI 182
 AAGRFGSGWAWLVVNNGE+EI STP QD+P+MEGK PILGLDVWEHAYYLKYQN+RP+YI
 Sbjct: 123 AAGRFGSGWAWLVVNNGEIEIMSTPIQDNPLMEGKKPILGLDVWEHAYYLKYQNKRPDI 182

Query: 183 AAFWNIVNWDEVAKRYSEA 201
 +AFWN+VNWDEVA +YS+A
 Sbjct: 183 SAFWNVNVWDEVAAQYSQA 201

 >gi|23099387|ref|NP_692853.1| manganese superoxide dismutase [Oceanobacillus ihe
 gi|22777616|dbj|BAC13888.1| manganese superoxide dismutase [Oceanobacillus iheyen
 Length = 203

Score = 302 bits (773), Expect = 3e-81

Identities = 156/200 (78%), Positives = 169/200 (84%)

Query: 3 FELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKXXXXXXXXXXXX 62
 FELP LPY YDALEP IDKETMNIHHTKHHNTYVT LN ALEGH DLQ+K
 Sbjct: 4 FELPELPYAYDALEPTIDKETMNIHHTKHHNTYVTKLNDALEGHADLQSKSVEELISNLD 63

Query: 63 XXXXIIRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFKAA 122
 +TAVRNNGGGHANHSLFW +LSPNGGGEPTGELA+ IN KFGS FK+EF+ A
 Sbjct: 64 AVPENAKTAVRNNGGGHANHSLFWKLLSPNGGGEPTGELADKINAKFGSLDKFKEEFAAA 123

Query: 123 AAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRREPEYI 182
 AAGRFGSGWAWLVVNNGELEITSTPNQDSP+MEGKTP+LGLDVWEHAYYLKYQN+RPEYI
 Sbjct: 124 AAGRFGSGWAWLVVNNGELEITSTPNQDSPLMGKTPVLGLDVWEHAYYLKYQNKREPEYI 183

Query: 183 AAFWNIVNWDEVAKRYSEAK 202
 +AFWN+VNWD+VAK Y EAK



NCBI

NCBI Conserved Domain Search

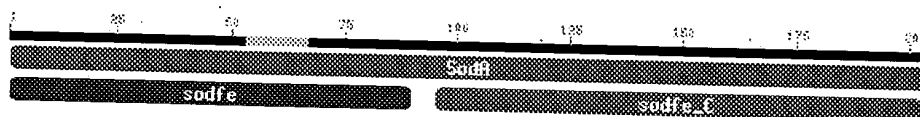
[New Search](#) [PubMed](#) [Nucleotide](#) [Protein](#) [Structure](#) [CDD](#) [Taxonomy](#) [Help?](#)

RPS-BLAST 2.2.6 [Apr-09-2003]

Query= local sequence:
(204 letters)

Database: #cdd.v1.62
11,088 PSSMs; 2,717,223 total columns

Click on boxes for multiple alignments



Domain Relatives

- .. This CD alignment includes 3D structure. To display structure, download [Cn3D!](#)

PSSMs producing significant alignments:

Score E
(bits) value

- gnl|CDD|10475 COG0605, SodA, Superoxide dismutase [Inorganic ion transport a... 282 3e-77
- gnl|CDD|17208 pfam02777, sodfe_C, Iron/manganese superoxide dismutases, C-te... 181 7e-47
- gnl|CDD|15113 pfam00081, sodfe, Iron/manganese superoxide dismutases, alpha-... 106 1e-24

- gnl|CDD|10475, COG0605, SodA, Superoxide dismutase [Inorganic ion transport and metabolism]

CD-Length = 204 residues, 99.5% aligned
Score = 282 bits (722), Expect = 3e-77

Query: 1	MPFELPALPYPYDALEPHIDKETMNIHHTKHNTYVTNLNAALEGHPDLQNKSLLELLSN	60
Sbjct: 2	MAYELPELPYAYDALEPHISAETMELHHDKHHQTYVNNLNAALEGLTE---ELEDLSLEE	58
Query: 61	LEALPESIRTAVRNGGGHANHSLFWTILSPN-GGGEPTGELAEAINKKFGSFTAFKDEF	119
Sbjct: 59	I I KKL AGLPAALFNNAGGHWNHSLFWENLSPGGGGGKPTGELAAAINKDFGSFDKFKEEF	118
Query: 120	SKAAAGRFGSGWAWLVVN-NGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNR	178
Sbjct: 119	TAAAASVFGSGWAWLVYDPDGKLEIVSTYNQDTPLMWGSPVLLGLDVWEHAYYLDYGNRR	178
Query: 179	PEYIAAFWNIVNWDEVAKRYSEAKAK	204
Sbjct: 179	PDYVEAFWNVNVWDEVBERFEAAKKE	204

- gnl|CDD|17208, pfam02777, sodfe_C, Iron/manganese superoxide dismutases, C-terminal domain. superoxide dismutases (SODs) catalyse the conversion of superoxide radicals to molecular oxygen. Three evolutionarily distinct families of SODs are known, of which the Mn/Fe-binding family is one. In humans, there is a cytoplasmic Cu/Zn SOD, and a mitochondrial Mn/Fe SOD. C-terminal domain is a mixed alpha/beta fold.

CD-Length = 111 residues, 100.0% aligned
Score = 181 bits (460), Expect = 7e-47

```
Query: 95  GEPTGELAEAINKKFGSFTAFKDEFSKAAAGRFGSGWAWLVVNN--GELEITSTPNQDSP 152
Sbjct: 1   GEPTGELADAIDEDFGSFEKFKKEEFTAAAAGVFGSGWAWLVYDNEGKKLAIVSTPNQDNF 60

Query: 153 IMECKTPILGLDVWEHAYYLKYQNRPEYIAAFWNIVNWDEVAKRYSEAKA 203
Sbjct: 61  LTEGLTPLLGLDVWEHAYYLDYQNRREFDYLKAFWNVNVWDEVSKRFEAAARK 111
```

- [gnl|CDD|15113](#), pfam00081, sodfe, Iron/manganese superoxide dismutases, alpha-hairpin domain. superoxide dismutases (SODs) catalyse the conversion of superoxide radicals to molecular oxygen. Three evolutionarily distinct families of SODs are known, of which the Mn/Fe-binding family is one. In humans, there is a cytoplasmic Cu/Zn SOD, and a mitochondrial Mn/Fe SOD. N-terminal domain is a long alpha antiparallel hairpin. A small fragment of YTRE_LEPBI matches well - sequencing error?

CD-Length = 83 residues, 100.0% aligned
Score = 106 bits (267), Expect = 1e-24

```
Query: 1  MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKSLLEELLSN 60
Sbjct: 1  MKYELPPLPYDYDALEPHISKETLEFHGKHHQTYVNNLNAALEGLEEARKE-----L 53

Query: 61  LEALPESIRTAVRNNGGGHANHSLFWTILS 90
Sbjct: 54  EELIIAALQGALRNNGGGHWNHTLEFWKNLS 83
```

[Help](#) | [Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLN](#) | [NIH](#)



results of BLAST

BLASTP 2.2.6 [Apr-09-2003]

RID: 1059166006-9282-7710.BLASTQ1

Query=

(622 letters)

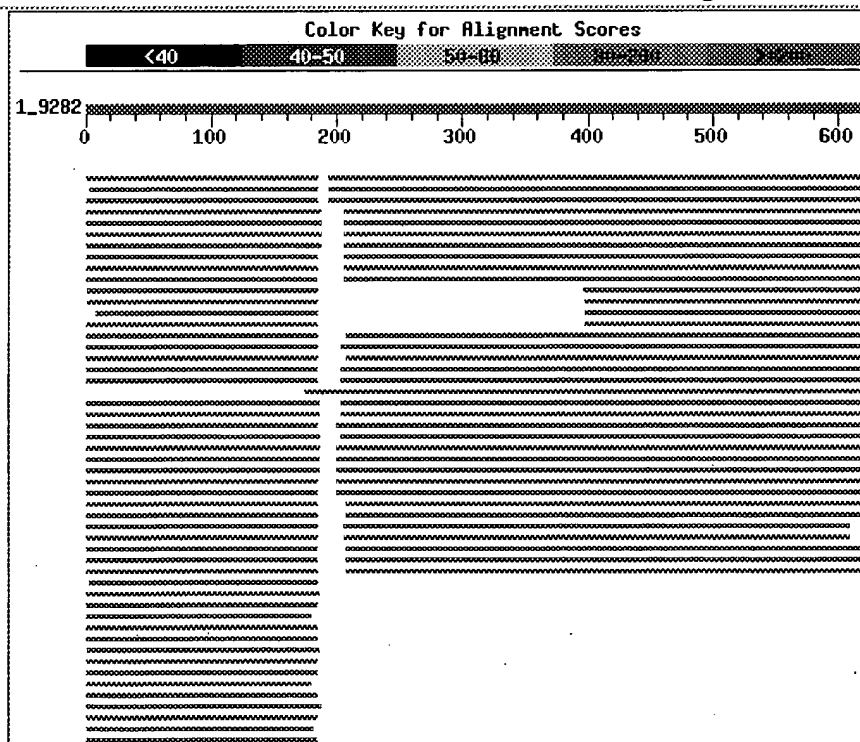
Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF

1,475,854 sequences; 475,827,238 total letters

[Taxonomy reports](#)

Distribution of 100 Blast Hits on the Query Sequence

Mouse-over to show defline and scores. Click to show alignments



Related Structures

Sequences producing significant alignments:

Score E
(bits) Value

gi	529984	gb	AAA23210.1	neurotoxin type F	>gi	971349	emb	C...	756	0.0
gi	3805790	emb	CAA73972.1	bonT [Clostridium botulinum]					654	0.0
gi	399137	sp	P30996	BXF CLOBO Botulinum neurotoxin type F p...					645	0.0
gi	479333	pir	S33411	botulinum neurotoxin type F - Clostri...					593	e-168
gi	9650999	dbj	EAB03512.1	type E botulinum toxin [Clostrid...					565	e-160
gi	98569	pir	S21178	botulinum neurotoxin type E precursor ...					564	e-159
gi	399135	sp	Q00496	BXE CLOBO Botulinum neurotoxin type E p...					563	e-159
gi	399136	sp	P30995	BXE CLOEU Botulinum neurotoxin type E p...					559	e-158

gi	9955895	dbj	EAB12249.1	type E botulinum toxin [Clostrid...	559	e-158
gi	21832314	dbj	BAC05434.1	type E botulinum toxin [Clostri...	559	e-158
gi	1084233	pir	S48110	neurotoxin type F - Clostridium botu...	392	e-107
gi	134649	sp	P28760	SODM BACCA Superoxide dismutase [Mn] >g...	387	e-106
gi	134650	sp	P00449	SODM BACST Superoxide dismutase [Mn] >g...	386	e-106
gi	7798571	dbj	BAA95631.1	superoxide dismutase [Geobacillu...	382	e-105
gi	448837	prf	1918164A	superoxide dismutase	381	e-104
gi	7438650	pir	S48109	neurotoxin type F - Clostridium botu...	330	3e-89
gi	15613972	ref	NP 242275.1	superoxide dismutase [Bacillus...	327	4e-88
gi	30264347	ref	NP 846724.1	superoxide dismutase, Mn [Baci...	318	2e-85
gi	3790114	emb	CAA05291.1	manganese superoxide dismutase [...]	315	1e-84
gi	3355882	dbj	BAA31974.1	superoxide dismutase [Bacillus s...	315	1e-84
gi	32468794	emb	CAB14432.2	superoxide dismutase [Bacillus ...]	313	5e-84
gi	16079557	ref	NP 390381.1	superoxide dismutase [Bacillus...	312	1e-83
gi	23099387	ref	NP 692853.1	manganese superoxide dismutase...	311	2e-83
gi	23200182	pdb	1JR9	Chain A, Crystal Structure Of Mangan...	310	4e-83
gi	30022355	ref	NP 833986.1	Superoxide dismutase [Mn] [Bac...	308	2e-82
gi	407787	emb	CAA50146.1	botulinum neurotoxin type E [Clos...	298	1e-79
gi	1084230	pir	S48106	neurotoxin type E - Clostridium botu...	297	4e-79
gi	16800546	ref	NP 470814.1	superoxide dismutase [Listeria...	293	6e-78
gi	2499920	sp	Q60393	BXG CLOBO Botulinum neurotoxin type G ...	292	1e-77
gi	29375094	ref	NP 814247.1	superoxide dismutase, Mn [Ente...	292	1e-77
gi	629166	pir	S39791	neurotoxin - Clostridium botulinum	291	2e-77
gi	15419708	gb	AAK97132.1	neurotoxin type B [Clostridium b...	291	2e-77
gi	15982937	gb	AAL11498.1	type B cryptic neurotoxin [Clost...	291	3e-77
gi	16803479	ref	NP 464964.1	superoxide dismutase [Listeria...	291	3e-77
gi	3805785	emb	CAA73968.1	bont [Clostridium botulinum]	290	4e-77
gi	7433329	pir	S20019	superoxide dismutase (EC 1.15.1.1) (...)	288	2e-76
gi	24137633	dbj	BAC22064.1	neurotoxin [Clostridium botulinum]	288	3e-76
gi	3320377	gb	AAC26483.1	manganese superoxide dismutase [V...	284	3e-75
gi	134666	sp	P28763	SODM LISIV Superoxide dismutase [Mn] >g...	284	3e-75
gi	399134	sp	P10844	BXB CLOBO Botulinum neurotoxin type B p...	280	4e-74
gi	11514111	pdb	1EPW	Chain A, Crystal Structure Of Clostr...	280	5e-74
gi	15599664	ref	NP 253158.1	superoxide dismutase [Pseudomo...	280	6e-74
gi	18251976	gb	AAL66183.1	AF464912 1 neurotoxin type A LHN ...	279	1e-73
gi	2118642	pir	I40631	non-proteolytic botulinum neurotoxin...	278	1e-73
gi	399133	sp	P10845	BXA1 CLOBO Botulinum neurotoxin type A ...	278	2e-73
gi	6137387	pdb	3BTA	Chain A, Crystal Structure Of Botulin...	278	2e-73
gi	279630	pir	BTCLAB	bontoxilysin (EC 3.4.24.69) A precurs...	278	2e-73
gi	40382	emb	CAA36289.1	type A neurotoxin (AA 1-1296) [Clo...	278	2e-73
gi	1084032	pir	A53294	superoxide dismutase (EC 1.15.1.1) (...)	275	1e-72
gi	23135202	ref	ZP 00116957.1	hypothetical protein [Cytoph...	274	3e-72
gi	32492801	gb	AAP85516.1	SodA [Aeromonas salmonicida subs...	274	4e-72
gi	21231716	ref	NP 637633.1	superoxide dismutase [Xanthomo...	272	1e-71
gi	21243120	ref	NP 642702.1	superoxidase dismutase [Xantho...	271	2e-71
gi	1150790	gb	AAB47971.1	superoxide dismutase [Xanthomonas...	271	2e-71
gi	31074373	gb	AAP41921.1	Mn/Fe superoxide dismutase [Chlo...	270	4e-71
gi	26987682	ref	NP 743107.1	superoxide dismutase (Mn) [Pse...	270	4e-71
gi	18378721	gb	AAL68691.1	AF462457 1 superoxide dismutase [...]	269	1e-70
gi	2500831	sp	P77929	SODM PSEPU Superoxide dismutase [Mn] >...	268	3e-70
gi	8977980	emb	CAB95744.1	superoxide dismutase [Staphyloco...	267	3e-70
gi	1711456	sp	P09214	SODM THETH Superoxide dismutase [Mn] >...	266	6e-70
gi	27468158	ref	NP 764795.1	superoxide dismutase SodA [Sta...	266	8e-70
gi	77595	pir	S07147	superoxide dismutase (EC 1.15.1.1) (Mn...	266	9e-70
gi	15924543	ref	NP 372077.1	superoxide dismutase [Staphylo...	266	9e-70
gi	23126478	ref	ZP 00108372.1	hypothetical protein [Nostoc...	266	1e-69
gi	23470984	ref	ZP 00126316.1	hypothetical protein [Pseudo...	264	3e-69
gi	17227566	ref	NP 484114.1	superoxide dismutase [Nostoc s...	263	7e-69
gi	23061777	ref	ZP 00086602.1	hypothetical protein [Pseudo...	263	7e-69
gi	20137335	sp	Q45894	BXA2 CLOBO Botulinum neurotoxin type ...	262	1e-68
gi	11121341	emb	CAC14833.1	superoxide dismutase [Staphyloc...	261	3e-68
gi	15806297	ref	NP 295003.1	superoxide dismutase (sodA), M...	261	3e-68
gi	134640	sp	P23744	SODF METJ Superoxide dismutase [Mn-Fe] ...	260	4e-68
gi	1711455	sp	P53653	SODM THEAQ Superoxide dismutase [Mn] >...	259	1e-67
gi	23200074	pdb	1GV3	Chain A, The 2.0 Angstrom Resolution...	258	2e-67

gi	1711423	sp	P50059	SOD2 PLEBO	Superoxide dismutase [Mn] 2...	258	2e-67
gi	32474407	ref	NP_867401.1		superoxide dismutase, Mn famil...	257	4e-67
gi	2147809	pir	S70582	botulinum neurotoxin type Dsa precu...		254	3e-66
gi	1711438	sp	P53642	SODM BORPE	Superoxide dismutase [Mn] >...	253	6e-66
gi	6939795	dbj	EAA90661.1		neurotoxin [Clostridium botulinum]	252	1e-65
gi	15022165	gb	AAK72964.2		tetanus toxin [Clostridium tetani]	252	2e-65
gi	22995080	ref	ZP_00039563.1		hypothetical protein [Xylella...	251	2e-65
gi	28373188	ref	NP_783831.1		tetanus toxin tetX [Clostridiu...	251	2e-65
gi	22995085	ref	ZP_00040357.1		hypothetical protein [Xylella...	250	4e-65
gi	260239	gb	AAB24244.1		type D neurotoxin [Bacteriophage d...	250	4e-65
gi	24182472	gb	AAN16456.2		manganese-containing superoxide ...	250	6e-65
gi	32035144	ref	ZP_00135190.1		hypothetical protein [Actino...	250	7e-65
gi	1711422	sp	P50058	SOD1 PLEBO	Superoxide dismutase [Mn] 1...	249	1e-64
gi	4579738	dbj	BAA75084.1		NTX [Clostridium botulinum D phage]	249	1e-64
gi	115188	sp	P19321	BXD CLOBO	Botulinum neurotoxin type D p...	248	2e-64
gi	28199854	ref	NP_780168.1		superoxide dismutase [Xylella ...	246	6e-64
gi	15839203	ref	NP_299891.1		superoxide dismutase [MN] [Xyl...	244	2e-63
gi	16124175	ref	NP_407488.1		superoxide dismutase [Mn] [Yer...	243	7e-63
gi	30023475	ref	NP_835106.1		Superoxide dismutase [Mn] [Bac...	243	9e-63
gi	541376	pir	PN0614		superoxide dismutase (EC 1.15.1.1) (M...	243	1e-62
gi	1711458	sp	P53655	SODM YEREN	Superoxide dismutase [Mn] >...	243	1e-62
gi	23040106	ref	ZP_00071638.1		hypothetical protein [Tricho...	242	1e-62
gi	15804498	ref	NP_290538.1		superoxide dismutase, manganes...	242	1e-62
gi	15551671	emb	CAC69393.1		manganese superoxide dismutase...	242	1e-62
gi	24115202	ref	NP_709712.1		superoxide dismutase, manganes...	241	2e-62
gi	16329169	gb	AAG44813.2	AF273269.1	superoxide dismutase [...	241	2e-62
gi	16131748	ref	NP_418344.1		superoxide dismutase, manganes...	241	2e-62

Alignments

Get selected sequences

Select all

Deselect all

☐ >gi|529984|gb|AAA23210.1| neurotoxin type F
gi|971349|emb|CAA57358.1| BoNT/F [Clostridium botulinum]
Length = 1278

Score = 756 bits (1953), Expect = 0.0

Identities = 403/429 (93%), Positives = 406/429 (94%), Gaps = 1/429 (0%)

Query: 194 LVPRGPGSKAPPRLCIRVNNRELFFVA SESSYNENDINTPKEIDDTTXXXXXXXXXXDEV 253
++PR G+KAPPRLCIRVNNRELFFVA SESSYNENDINTPKEIDDTT DEV
Sbjct: 432 VIPR-KGTKAPPRLCIRVNNRELFFVA SESSYNENDINTPKEIDDTTNLNNNYRNNLDEV 490

Query: 254 IL DYNSETIPQISNQT LNTLVQDDSYVPRYDSNGTSEIEEHNVDLNVFFYLHAQKVPEG 313
IL DYNSETIPQISNQT LNTLVQDDSYVPRYDSNGTSEIEEHNVDLNVFFYLHAQKVPEG
Sbjct: 491 IL DYNSETIPQISNQT LNTLVQDDSYVPRYDSNGTSEIEEHNVDLNVFFYLHAQKVPEG 550

Query: 314 ETNISLTSSIDTAL SEESQVYTFFSSEFINTINKPVHAALFISWINQVIRDFTTTEATQKS 373
ETNISLTSSIDTAL SEESQVYTFFSSEFINTINKPVHAALFISWINQVIRDFTTTEATQKS
Sbjct: 551 ETNISLTSSIDTAL SEESQVYTFFSSEFINTINKPVHAALFISWINQVIRDFTTTEATQKS 610

Query: 374 TFDKIADISLVVPYVGLALNIGNEVQKENFKEAFELLGAGILLEFVPELLIPTILVFTIK 433
TFDKIADISLVVPYVGLALNIGNEVQKENFKEAFELLGAGILLEFVPELLIPTILVFTIK
Sbjct: 611 TFDKIADISLVVPYVGLALNIGNEVQKENFKEAFELLGAGILLEFVPELLIPTILVFTIK 670

Query: 434 SFIGSSEXXXXXXXXXXXSLMERETKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQ 493
SFIGSSE SLMERETKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQ
Sbjct: 671 SFIGSSENKNKIIKAINNSLMERETKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQ 730

Query: 494 VDAIKTVIEYKNNYTS DERNRLESEYNINNIREELNKKVSLAMENIERFITESSIFYLM 553
VDAIKTVIEYKNNYTS DERNRLESEYNINNIREELNKKVSLAMENIERFITESSIFYLM
Sbjct: 731 VDAIKTVIEYKNNYTS DERNRLESEYNINNIREELNKKVSLAMENIERFITESSIFYLM 790

Query: 554 KLINEAKVSKLREYDEGVKEYLLDYISEHRSILGNSVQELNDLVTSTLNNSIPFELSSYT 613
 KLINEAKVSKLREYDEGVKEYLLDYISEHRSILGNSVQELNDLVTSTLNNSIPFELSSYT
 Sbjct: 791 KLINEAKVSKLREYDEGVKEYLLDYISEHRSILGNSVQELNDLVTSTLNNSIPFELSSYT 850

Query: 614 NDKILILYF 622
 NDKILILYF
 Sbjct: 851 NDKILILYF 859

☐ >gi|3805790|emb|CAA73972.1| bonT [Clostridium botulinum]
 Length = 1280

Score = 654 bits (1686), Expect = 0.0
 Identities = 347/429 (80%), Positives = 376/429 (87%), Gaps = 1/429 (0%)

Query: 194 LVPRGPGSKAPPRLCIRVNNRELFFVASESSYNENDINTPKEIDDTTXXXXXXXXXXDEV 253
 ++PR G+K P LCIRVNNRELFFVASESSYNE+DINTPKEIDDTT DEV
 Sbjct: 432 IIPR-KGTKQSPSLCIRVNNRELFFVASESSYNESDINTPKEIDDTTNLNNNYRNNLDEV 490

Query: 254 ILDYNSETIPQISNQTTLNTLVQDDSYVPYDSNGTSEIEEHNVDLNVFFYLHAQKVPEG 313
 ILDYNSETIPQISN+TLNTLVQD+SYVPYDSNGTSEIEE++VVD NVFFYLHAQKVPEG
 Sbjct: 491 ILDYNSETIPQISNRTLNTLVQDNSYVPYDSNGTSEIEEYDVDFNVFFYLHAQKVPEG 550

Query: 314 ETNISLTSSIDTALSEESQVYTFSSSEFINTINKPVHAALFISWINQVIRDFTEATQKS 373
 ETNISLTSSIDTAL EES+VYTFSSSEFI+TINKPV+AALFI WI++VIRDFTEATQKS
 Sbjct: 551 ETNISLTSSIDTALLEESKVYTFSSSEFIDTINKPVNAALFIDWISKVIRDFTEATQKS 610

Query: 374 TFDKIADISLVVPYVGLALNIGNEVQKENFKEAFELLGAGILLEFVPELLIPTILVFTIK 433
 TDKIADISL+VPYVGLALNI E +K NF+EAFELLGAGILLEFVPEL IP ILVFTIK
 Sbjct: 611 TVDKIADISLVVPYVGLALNIVIEAEKGNFEEAFELLGAGILLEFVPELTIPVILVFTIK 670

Query: 434 SFIGSSEXXXXXXXXXXXXXSLMERETKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQ 493
 S+I S E SL+ERE KWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQ
 Sbjct: 671 SYIDSYENKNKAIKAINNSLIEREAKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQ 730

Query: 494 VDAIKTVIEYKYNNTSDERNRLESEYNINNIREELNKKVSLAMENIERFITESSIFYLM 553
 VDAIKT IEYKYNNTSDE+NRLES+YNINNI EELNKKVSLAM+NIERF+TESSI YLM
 Sbjct: 731 VDAIKTAIEYKYNNTSDEKNRLESKYNINNIEEELNKKVSLAMKNIERFMTESSISYLM 790

Query: 554 KLINEAKVSKLREYDEGVKEYLLDYISEHRSILGNSVQELNDLVTSTLNNSIPFELSSYT 613
 KLINEA+V KL+EYD+ VK LLDYI H+ ILG +EL DLVTSTLN+SIPFELSSYT
 Sbjct: 791 KLINEAEVGKLKEYDKHVKSDDLIDYILYHKLILGEQTKELIDLVTSTLNSSIPFELSSYT 850

Query: 614 NDKILILYF 622
 NDKILI+YF
 Sbjct: 851 NDKILIIYF 859

☐ >gi|399137|sp|P30996|BXF_CLOBO Botulinum neurotoxin type F precursor (BoNT/F) (B
 gi|2118643|pir|I40813 neurotoxin type F - Clostridium botulinum
 gi|144867|gb|AAA23263.1| neurotoxin type F
 gi|383771|prf|1904210A neurotoxin F
 Length = 1274

Score = 645 bits (1665), Expect = 0.0
 Identities = 347/429 (80%), Positives = 375/429 (87%), Gaps = 2/429 (0%)

Query: 194 LVPRGPGSKAPPRLCIRVNNRELFFVASESSYNENDINTPKEIDDTTXXXXXXXXXXDEV 253
 ++PR G+KAPPRLCIRVNN ELFFVASESSYNENDINTPKEIDDTT DEV
 Sbjct: 432 VIPR-KGTKAPPRLCIRVNNSELFFVASESSYNENDINTPKEIDDTTNLNNNYRNNLDEV 490

Query: 254 ILDYNSETIPQISNQTTLNTLVQDDSYVPYDSNGTSEIEEHNVDLNVFFYLHAQKVPEG 313
 ILDYN+TIPQISN+TLNTLVQD+SYVPYDSNGTSEIEE++VVD NVFFYLHAQKVPEG
 Sbjct: 491 ILDYNSETIPQISNRTLNTLVQDNSYVPYDSNGTSEIEEYDVDFNVFFYLHAQKVPEG 550

Query: 314 ETNISLTSSIDTALSEESQVYTFSSSEFINTINKPVHAALFISWINQVIRDFTEATQKS 373
 ETNISLTSSIDTAL EES+ FFSSEFI+TINKPV+AALFI WI++VIRDFTEATQKS

Sbjct: 551 ETNISLTSSIDTALLEESK-DIFFSSEFIDTINKPVNAALFIDWISKVIRDFTEATQKS 609

Query: 374 TFDKIADISLVVPYVGLALNIGNEVQKENFKEAFELLGAGILLEFVPELLIPTILVFTIK 433
T DKIADISL+VPYVGLALNI E +K NF+EAFELLG GILLEFVPEL IP ILVFTIK

Sbjct: 610 TVDKIADISLIVPYVGLALNIIIEAEKGNFEEAFELLGVGILLEFVPELTIPVILVFTIK 669

Query: 434 SFIGSSEXXXXXXXXXXXSLMERETKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQ 493
S+I S E SL+ERE KWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQ

Sbjct: 670 SYIDSYENKNKAIKAINNSLIEREAKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQ 729

Query: 494 VDAIKTVIEYKYNNTSDERNRLESEYNINNIREELNKKVSLAMENIERFITESSIFYLM 553
VDAIKT IEYKYNNTSDE+NRLESEYNINNI EELNKKVSLAM+NIERF+TESSI YLM

Sbjct: 730 VDAIKTAIEYKYNNTSDEKNRLESEYNINNIEEELNKKVSLAMKNIERFMTESSISYLM 789

Query: 554 KLINEAKVSKLREYDEGVKEYLLDYISEHRSILGNSVQELNDLVTSTLNNSIPFELSSYT 613
KLINEAKV KL++YD VK LL+YI +HRSILG EL+DLVTSTLN+SIPFELSSYT

Sbjct: 790 KLINEAKVGKLLKYDNHVKSDLLNYILDHRSILGEQTNELSDLVTSTLNNSIPFELSSYT 849

Query: 614 NDKILILYF 622
NDKILI+YF

Sbjct: 850 NDKILIIYF 858

☐ >gi|479333|pir|S33411 botulinum neurotoxin type F - Clostridium barati
gi|49139|emb|CAA48329.1 neurotoxin type F [Clostridium baratii]
Length = 1268

Score = 593 bits (1530), Expect = e-168

Identities = 314/414 (75%), Positives = 355/414 (85%), Gaps = 1/414 (0%)

Query: 207 LCIRVNNRELFFVASESSYNENDINTPKEIDDTXXXXXXXXXXDEVILDYNSETIPQIS 266
LCI+VNNR+LFFVASESSYNEN IN+PKEIDDT DEVILDYNS+ IP +S

Sbjct: 434 LCIKVNNRDLFFVASESSYNENGINSPEIDDTTITNNNYKKNLDEVILDYNSDAIPNLS 493

Query: 267 NQTLNLTIVQDDSYVPRYDSNGTSEIEHNVVDLNVFFYLHAQKVPEGETNISLTSSIDTA 326
++ LNT Q+DSYVP+YDSNGTSEI+E+ V LNVFFYL+AQK PEGE+ ISLTSS++TA

Sbjct: 494 SRLLNTTAQNDSYVPKYDSNGTSEIKEYTVDKLNVFFYLYAQKAPEGESAISLTSSVNTA 553

Query: 327 LSEESQVYTFSSSEFINTINKPVHAALFISWINQVIRDFTEATQKSTFDKIADISLVVP 386
L + S+VYTFSS+FINT+NKPV AALFISWI QVI DFTTEATQKST DKIADISL+VP

Sbjct: 554 LLDASKVYTFSSDFINTVNKPVQAALFISWIQQVINDFTTEATQKSTIDKIADISLIVP 613

Query: 387 YVGLALNIGNEVQKENFKEAFELLGAGILLEFVPELLIPTILVFTIKSFIGSSEXXXXXX 446
YVGLALNIGNEVQK NFKEA ELLGAGILLEFVPELLIPTILVFTIKSFI S +

Sbjct: 614 YVGLALNIGNEVQKGNFKEAIELLGAILEFVPELLIPTILVFTIKSFINSDDSKNKII 673

Query: 447 XXXXXSLMERETKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQVDAIKTVIEYKYN 506
+L ERE KWKE+YSWIVSNWLTRINTQFNKRKEQMYQALQNQVD IK +IEYKYN

Sbjct: 674 KAINNALRERELKWKEVYSWIVSNWLTRINTQFNKRKEQMYQALQNQVDGIKKIIEYKYN 733

Query: 507 NYTSDERNRLESEYNINNIREELNKKVSLAMENIERFITESSIFYLMKLINEAKVSKLRE 566
NYT DE+NRL +EYNI +I+EELNKKVSLAM+NI+RF+TESSI YLMKLINEAK++KL E

Sbjct: 734 NYTLDEKNRLRAEYNIYSIKEELNKKVSLAMQNIDRFLTESSISYLMKLINEAKINKLSE 793

Query: 567 YDEGVKEYLLDYISEHRSILG-NSVQELNDLVTSTLNNSIPFELSSYTNDKILI 619
YD+ V +YLL+YI E+ S LG +SV ELN+LV++TLNNSIPFELS YTNDKILI

Sbjct: 794 YDKRVNQYLLNYILENSSTLGTSSVPELNNLVSNLTLNNSIPFELSEYTNDKILI 847

☐ >gi|9650999|dbj|BAB03512.1 type E botulinum toxin [Clostridium butyricum]

gi	9651001	dbj	BAB03513.1	type E botulinum toxin [Clostridium butyricum]
gi	9651003	dbj	BAB03514.1	type E botulinum toxin [Clostridium butyricum]
gi	9651005	dbj	BAB03515.1	type E botulinum toxin [Clostridium butyricum]
gi	9651007	dbj	BAB03516.1	type E botulinum toxin [Clostridium butyricum]
gi	9651009	dbj	BAB03517.1	type E botulinum toxin [Clostridium butyricum]
gi	9651011	dbj	BAB03518.1	type E botulinum toxin [Clostridium butyricum]
gi	9651013	dbj	BAB03519.1	type E botulinum toxin [Clostridium butyricum]

gi	9651015	dbj	BAB03520.1	type E botulinum toxin [Clostridium butyricum]
gi	9651017	dbj	BAB03521.1	type E botulinum toxin [Clostridium butyricum]
gi	9651019	dbj	BAB03522.1	type E botulinum toxin [Clostridium butyricum]

Length = 1251

Score = 565 bits (1457), Expect = e-160
Identities = 289/416 (69%), Positives = 350/416 (84%)

Query: 207 LCIRVNNRELFFVASESSYNENDINTPKEIDDTTXXXXXXXXXXDEVILDYNSETIPQIS 266
+CI +NN ELFFVASE+SYN+++INTPKEIDDT D+VIL++NSE+ P +S

Sbjct: 425 ICIEINNGELFFVASENSYNDDNINTPKEIDDTVTSNNNYENDLDQVILNFNSESAPGLS 484

Query: 267 NQTLNLTIVQDDSYVPRYDSNGTSEIEHNVDLNVFFYLHAQKVPEGETNISLTSSIDTA 326
++ LN +Q+D+Y+P+YDSNGTS+IE+H+V +LNVFFYL AQKVPEGE N++LTSSIDTA

Sbjct: 485 DEKLNLTIQNDAYIPKYDSNGTSDIEQHDVNELNVFFYLDAQKVPEGENNVNLTSSIDTA 544

Query: 327 LSEESQVYTFFSSEFINTINKPVHAALFISWINQVIRDFTTTEATQKSTFDKIADISLVVP 386
L E+ ++YTFFSSEFIN +NKPV AALF+SWI QV+ DFTTEA QKST DKIADIS+VVP

Sbjct: 545 LLEQPKIYTFFSSEFINNVNKPVQAALFVSWIQVLVDFTTTEANQKSTVDKIADISIVVP 604

Query: 387 YVGLALNIGNEVQKENFKEAFELLGAGILLEFVPELLIPTILVFTIKSFIGSSEXXXXXX 446
Y+GLALNIGNE QK NFK+A ELLGAGILLEFVPELLIPTILVFTIKSF+GSS+

Sbjct: 605 YIGLALNIGNEAQKGNFKDALELLGAGILLEFVPELLIPTILVFTIKSFLGSSDNKNKVI 664

Query: 447 XXXXXSLMERETKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQVDAIKTVIEYKYN 506
+L ER+ KWKE+YS+IVSNW+T+INTQFNKRKEQMYQALQNQV+A+KT+IE+KYN

Sbjct: 665 KAINNALKERDEKWKEVYSFIVSNWMTKINTQFNKRKEQMYQALQNQVNAIKTIIIEFKYN 724

Query: 507 NYTSDERNRLESEYNINNIREELNKKVSLAMENIERFITESSIFYLMKLINEAKVSKLRE 566
+YT +E+ L++ Y+I I ELN+KVS+AM NI+RF+TESSI YLMKLINE K++KLRE

Sbjct: 725 SYTLEEKKEKKNYDIEQIENELNQKVSIAMNNIDRFLTESSISYLMKLINEVKINKLRE 784

Query: 567 YDEGVKEYLLDYISEHRSILGNSVQELNDLVTSTLNNSIPFELSSYTNDKILILYF 622
YDE VK YLLDYI +H SILG S QELN +V TLNNSIPF+LSSYT+DKILI YF

Sbjct: 785 YDENVKTYLLDYIIQHGSILGESQQELNSMVIDTLNNSIPFKLSSYTDDKILISYF 840

>gi	98569	pir	S21178	botulinum neurotoxin type E precursor - Clostridium botuli
gi	40398	emb	CAA44558.1	botulinum neurotoxin type E [Clostridium botulinum]
gi	19744173	dbj	BAB86845.1	botulinum neurotoxin type E [Clostridium botulinum]

Length = 1252

Score = 564 bits (1453), Expect = e-159
Identities = 290/416 (69%), Positives = 350/416 (84%)

Query: 207 LCIRVNNRELFFVASESSYNENDINTPKEIDDTTXXXXXXXXXXDEVILDYNSETIPQIS 266
+CI +NN ELFFVASE+SYN+++INTPKEIDDT D+VIL++NSE+ P +S

Sbjct: 425 ICIEINNGELFFVASENSYNDDNINTPKEIDDTVTSNNNYENDLDQVILNFNSESAPGLS 484

Query: 267 NQTLNLTIVQDDSYVPRYDSNGTSEIEHNVDLNVFFYLHAQKVPEGETNISLTSSIDTA 326
++ LN +Q+D+Y+P+YDSNGTS+IE+H+V +LNVFFYL AQKVPEGE N++LTSSIDTA

Sbjct: 485 DEKLNLTIQNDAYIPKYDSNGTSDIEQHDVNELNVFFYLDAQKVPEGENNVNLTSSIDTA 544

Query: 327 LSEESQVYTFFSSEFINTINKPVHAALFISWINQVIRDFTTTEATQKSTFDKIADISLVVP 386
L E+ ++YTFFSSEFIN +NKPV AALF+SWI QV+ DFTTEA QKST DKIADIS+VVP

Sbjct: 545 LLEQPKIYTFFSSEFINNVNKPVQAALFVSWIQVLVDFTTTEANQKSTVDKIADISIVVP 604

Query: 387 YVGLALNIGNEVQKENFKEAFELLGAGILLEFVPELLIPTILVFTIKSFIGSSEXXXXXX 446
Y+GLALNIGNE QK NFK+A ELLGAGILLEF PELLIPTILVFTIKSF+GSS+

Sbjct: 605 YIGLALNIGNEAQKGNFKDALELLGAGILLEFEPELLIPTILVFTIKSFLGSSDNKNKVI 664

Query: 447 XXXXXSLMERETKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQVDAIKTVIEYKYN 506
+L ER+ KWKE+YS+IVSNW+T+INTQFNKRKEQMYQALQNQV+AIKT+IE KYN

Sbjct: 665 KAINNALKERDEKWKEVYSFIVSNWMTKINTQFNKRKEQMYQALQNQVNAIKTIIIESKYN 724

Query: 507 NYTSDERNRLESEYNINNIREELNKKVSLAMENIERFITESSIFYLMKLINEAKVSKLRE 566

+YT +E+N L ++Y+I I ELN+KVS+AM NI+RF+TESSI YLMKLINE K++KLRE
 Sbjct: 725 SYTLEEKNELTNKYDIKQIENELNQKVSAMNNIDRFLTESSISYLMKLINEVKINKLRE 784

Query: 567 YDEGVKEYLLDYISEHRSILGNSVQELNDLVTSTLNN SIPFELSSYTNDKILILYF 622
 YDE VK YLL+YI +H SILG S QELN +VT TLNNSIPF+LSSYT+DKILI YF
 Sbjct: 785 YDENVKTYLLNYIIQHGSILGESQQELNSMVTDTLNN SIPFKLSSYTDDKILISYF 840

>gi|399135|sp|Q00496|EXE_CLOBO Botulinum neurotoxin type E precursor (BoNT/E) (B
 gi|40394|emb|CAA43999.1| botulinum neurotoxin type E [Clostridium botulinum]
 Length = 1251

Score = 563 bits (1451), Expect = e-159
 Identities = 289/416 (69%), Positives = 350/416 (84%)

Query: 207 LCIRVNNRELFFVASESSYNENDINTPKEIDDTTXXXXXXXXXXDEVILDYNSETIPQIS 266
 +CI +NN ELFFVASE+SYN+++INTPKEIDDT D+VIL++NSE+ P +S
 Sbjct: 425 ICIEINNGELFFVASENSYNDDNINTPKEIDDTVTSNNNYENDLDQVILNFNSESAPGLS 484

Query: 267 NQTLNLTQDDSYVPRYDSNGTSEIEHNVDLNVFFYLHAQKVPEGETNISLTSSIDTA 326
 ++ LN +Q+D+Y+P+YDSNGTS+IE+H+V +LNVFFYL AQKVPEGE N++LTSSIDTA
 Sbjct: 485 DEKLNLTIQNDAYIPKYDSNGTSDIEQHDVNELNVFFYLDAQKVPEGENNVNLTSSIDTA 544

Query: 327 LSEESQVYTFFSSEFINTINKPVHAALFISWINQVIRDFTEATQKSTFDKIADISLVVP 386
 L E+ ++YTFFSSEFIN +NKPV AALF+SWI QV+ DFTTEA QKST DKIADIS+VVP
 Sbjct: 545 LLEQPKIYTFFSSEFINNVNKPVQAALFVSWIQVQLVDFTEANQKSTVDKIADISIVVP 604

Query: 387 YVGLALNIGNEVQKENFKEAFELLGAGILLEFVPELLIPTILVFTIKSFIGSSEXXXXXX 446
 Y+GLALNIGNE QK NFK+A ELLGAGILLEF PELLIPTILVFTIKSF+GSS+
 Sbjct: 605 YIGLALNIGNEAQKGNFKDALELLGAGILLEFEPELLIPTILVFTIKSFLGSSDNKNKVI 664

Query: 447 XXXXXSLMERETKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQVDAIKTVIEYKYN 506
 +L ER+ KWKE+YS+IVSNW+T+INTQFNKRKEQMYQALQNQV+AIKT+IE KYN
 Sbjct: 665 KAINNALKERDEKWKEVYSFIVSNWMTKINTQFNKRKEQMYQALQNQVNAIKTIIESKYN 724

Query: 507 NYTSDERNRLESEYNINNIREELNKKVSLAMENIERFITESSIFYLMKLINEAKVSKLRE 566
 +YT +E+N L ++Y+I I ELN+KVS+AM NI+RF+TESSI YLMK+INE K++KLRE
 Sbjct: 725 SYTLEEKNELTNKYDIKQIENELNQKVSAMNNIDRFLTESSISYLMKIINEVKINKLRE 784

Query: 567 YDEGVKEYLLDYISEHRSILGNSVQELNDLVTSTLNN SIPFELSSYTNDKILILYF 622
 YDE VK YLL+YI +H SILG S QELN +VT TLNNSIPF+LSSYT+DKILI YF
 Sbjct: 785 YDENVKTYLLNYIIQHGSILGESQQELNSMVTDTLNN SIPFKLSSYTDDKILISYF 840

>gi|399136|sp|P30995|EXE_CLOBU Botulinum neurotoxin type E precursor (BoNT/E) (B
 gi|7438649|pir|JH0256| botulinum neurotoxin type E precursor - Clostridium butyri
 gi|40380|emb|CAA43998.1| botulinum neurotoxin [Clostridium butyricum]
 Length = 1251

Score = 559 bits (1440), Expect = e-158
 Identities = 287/416 (68%), Positives = 347/416 (83%)

Query: 207 LCIRVNNRELFFVASESSYNENDINTPKEIDDTTXXXXXXXXXXDEVILDYNSETIPQIS 266
 +CI +NN ELFFVASE+SYN+++INTPKEIDDT D+VIL++NSE+ P +S
 Sbjct: 425 ICIEINNGELFFVASENSYNDDNINTPKEIDDTVTSNNNYENDLDQVILNFNSESAPGLS 484

Query: 267 NQTLNLTQDDSYVPRYDSNGTSEIEHNVDLNVFFYLHAQKVPEGETNISLTSSIDTA 326
 ++ LN +Q+D+Y+P+YDSNGTS+IE+H+V +LNVFFYL AQKVPEGE N++LTSSIDTA
 Sbjct: 485 DEKLNLTIQNDAYIPKYDSNGTSDIEQHDVNELNVFFYLDAQKVPEGENNVNLTSSIDTA 544


Query: 327 LSEESQVYTFFSSEFINTINKPVHAALFISWINQVIRDFTEATQKSTFDKIADISLVVP 386
 L E+ ++YTFFSSEFIN +NKPV AALF+ WI QV+ DFTTEA QKST DKIADIS+VVP
 Sbjct: 545 LLEQPKIYTFFSSEFINNVNKPVQAALFVGWIQVQLVDFTEANQKSTVDKIADISIVVP 604

Query: 387 YVGLALNIGNEVQKENFKEAFELLGAGILLEFVPELLIPTILVFTIKSFIGSSEXXXXXX 446
 Y+GLALNIGNE QK NFK+A ELLGAGILLEF PELLIPTILVFTIKSF+GSS+
 Sbjct: 605 YIGLALNIGNEAQKGNFKDALELLGAGILLEFEPELLIPTILVFTIKSFLGSSDNKNKVI 664

Query: 447 XXXXXSLMERETKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQVDAIKTVIEYKYN 506
 +L ER+ KWKE+YS+IVSNW+T+INTQFNKRKEQMYQALQNQV+A+K +IE KYN
 Sbjct: 665 KAINNALKERDEKWKEVYSFIVSNWMTKINTQFNKRKEQMYQALQNQVNALKAIIESKYN 724

Query: 507 NYTSDERNRLESEYNINNIREELNKKVSLAMENIERFITESSIFYLMKLINEAKVSKLRE 566
 +YT +E+N L ++Y+I I ELN+KVS+AM NI+RF+TESSI YLMKLINE K++KLRE
 Sbjct: 725 SYTLEEKNELTNKYDIEQIENELNQKVSAMNNIDRFLTESSISYLMKLINEVKINKLRE 784

Query: 567 YDEGVKEYLLDYISEHRSILGNSVQELNDLVTSTLNN SIPFELSSYTNDKILILYF 622
 YDE VK YLLDYI +H SILG S QELN +V TLNNSIPF+LSSYT+DKILI YF
 Sbjct: 785 YDENVKTYLLDYIIKHGSILGESQQELNSMVIDTLNNSIPFKLSSYTDDKILISYF 840

 >gi|9955895|dbj|BAB12249.1| type E botulinum toxin [Clostridium butyricum]
 Length = 1255

Score = 559 bits (1440), Expect = e-158
 Identities = 287/416 (68%), Positives = 347/416 (83%)

Query: 207 LCIRVNNRELFFVASESSYNENDINTPKEIDDTXXXXXXXXXXDEVILDYNSETIPQIS 266
 +CI +NN ELFFVASE+SYN+++INTPKEIDDT D+VIL++NSE+ P +S
 Sbjct: 428 ICIEINNGELFFVASENSYNDDNINTPKEIDDTVTSNNNYENDLDQVILNFNSESAPGLS 487

Query: 267 NQTLNLTQDDSYVPRYDSNGTSEIEEHNVDLNVFFYLHAQKVPEGETNISLTSSIDTA 326
 ++ LN +Q+D+Y+P+YDSNGTS+IE+H+V +LNVFFYL AQKVPEGE N++LTSSIDTA
 Sbjct: 488 DEKLNLTIQNDAYIPKYDSNGTSDIEQHDVNELNVFFYLDAQKVPEGENNVNLTSSIDTA 547


Query: 327 LSEESQVYTFSSSEFINTINKPVHAALFISWINQVIRDFTTEATQKSTFDKIADISLVVP 386
 L E+ ++YTFFSSEFIN +NKPV AALF+ WI QV+ DFTTEA QKST DKIADIS+VVP
 Sbjct: 548 LLEQPKIYTFSSSEFINNVNKPVAALFVGWIQQVLVDFTTEANQKSTVDKIADISIVVP 607

Query: 387 YVGLALNIGNEVQKENFKEAFELLGAGILLEFVPELLIPTILVFTIKSFIGSSEXXXXXX 446
 Y+GLALNIGNE QK NFK+A ELLGAGILLEF PELLIPTILVFTIKSF+GSS+
 Sbjct: 608 YIGLALNIGNEAQKGNFKDALELLGAGILLEFEPELLIPTILVFTIKSFLGSSDNKNKVI 667

Query: 447 XXXXXSLMERETKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQVDAIKTVIEYKYN 506
 +L ER+ KWKE+YS+IVSNW+T+INTQFNKRKEQMYQALQNQV+A+K +IE KYN
 Sbjct: 668 KAINNALKERDEKWKEVYSFIVSNWMTKINTQFNKRKEQMYQALQNQVNALKAIIESKYN 727

Query: 507 NYTSDERNRLESEYNINNIREELNKKVSLAMENIERFITESSIFYLMKLINEAKVSKLRE 566
 +YT +E+N L ++Y+I I ELN+KVS+AM NI+RF+TESSI YLMKLINE K++KLRE
 Sbjct: 728 SYTLEEKNELTNKYDIEQIENELNQKVSAMNNIDRFLTESSISYLMKLINEVKINKLRE 787

Query: 567 YDEGVKEYLLDYISEHRSILGNSVQELNDLVTSTLNN SIPFELSSYTNDKILILYF 622
 YDE VK YLLDYI +H SILG S QELN +V TLNNSIPF+LSSYT+DKILI YF
 Sbjct: 788 YDENVKTYLLDYIIKHGSILGESQQELNSMVIDTLNNSIPFKLSSYTDDKILISYF 843

 >gi|21832314|dbj|EAC05434.1| type E botulinum toxin [Clostridium butyricum]
 Length = 1252

Score = 559 bits (1440), Expect = e-158
 Identities = 287/416 (68%), Positives = 347/416 (83%)

Query: 207 LCIRVNNRELFFVASESSYNENDINTPKEIDDTXXXXXXXXXXDEVILDYNSETIPQIS 266
 +CI +NN ELFFVASE+SYN+++INTPKEIDDT D+VIL++NSE+ P +S
 Sbjct: 425 ICIEINNGELFFVASENSYNDDNINTPKEIDDTVTSNNNYENDLDQVILNFNSESAPGLS 484

Query: 267 NQTLNLTQDDSYVPRYDSNGTSEIEEHNVDLNVFFYLHAQKVPEGETNISLTSSIDTA 326
 ++ LN +Q+D+Y+P+YDSNGTS+IE+H+V +LNVFFYL AQKVPEGE N++LTSSIDTA
 Sbjct: 485 DEKLNLTIQNDAYIPKYDSNGTSDIEQHDVNELNVFFYLDAQKVPEGENNVNLTSSIDTA 544

Query: 327 LSEESQVYTFSSSEFINTINKPVHAALFISWINQVIRDFTTEATQKSTFDKIADISLVVP 386
 L E+ ++YTFFSSEFIN +NKPV AALF+ WI QV+ DFTTEA QKST DKIADIS+VVP
 Sbjct: 545 LLEQPKIYTFSSSEFINNVNKPVAALFVGWIQQVLVDFTTEANQKSTVDKIADISIVVP 604

Query: 387 YVGLALNIGNEVQKENFKEAFELLGAGILLEFVPELLIPTILVFTIKSFIGSSEXXXXXX 446
 Y+GLALNIGNE QK NFK+A ELLGAGILLEF PELLIPTILVFTIKSF+GSS+
 Sbjct: 605 YIGLALNIGNEAQKGNFKDALELLGAGILLEFEPELLIPTILVFTIKSFLGSSDNKNKVI 664

Query: 447 XXXXXSLMERETKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQVDAIKTVIEYKYN 506
 +L ER+ KWKE+YS+IVSNW+T+INTQFNKRKEQMYQALQNQV+A+K +IE KYN
 Sbjct: 665 KAINNALKERDEKWKEVYSFIVSNWMTKINTQFNKRKEQMYQALQNQVNALKAIIESKYN 724

Query: 507 NYTSDERNRLESEYNINNIREELNKKVSLAMENIERFITESSIFYLMKLINEAKVSKLRE 566
 +YT +E+N L ++Y+I I ELN+KVS+AM NI+RF+TESSI YLMKLINE K++KLRE
 Sbjct: 725 SYTLEEKNELTNKYDIEQIENELNQVSIAMNNIDRFLTESSISYLMKLINEVKINKLRE 784

Query: 567 YDEGVKEYLLDYISEHRSILGNSVQELNDLVTSTLNN SIPFELSSYTNDKILILYF 622
 YDE VK YLLDYI +H SILG S QELN +V TLNNSIPF+LSSYT+DKILI YF
 Sbjct: 785 YDENVKTYLLDYIIKHGSILGESQQELNSMVIDTLNNSIPFKLSSYTDDKILISYF 840

>gi|1084233|pir|S48110 neurotoxin type F - Clostridium botulinum (fragment)
 gi|407793|emb|CAA50152.1| botulinum neurotoxin type F [Clostridium botulinum]
 Length = 366

Score = 392 bits (1006), Expect = e-107
 Identities = 214/225 (95%), Positives = 214/225 (95%)

Query: 398 VQKENFKEAFELLGAGILLEFVPELLIPTILVFTIKSFIGSSEXXXXXXXXXXXXXSLMERE 457
 VQKENFKEAFELLGAGILLEFVPELLIPTILVFTIKSFIGSSE SLMERE
 Sbjct: 1 VQKENFKEAFELLGAGILLEFVPELLIPTILVFTIKSFIGSSENKNKI IKAINNLSLMERE 60

Query: 458 TKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQVDAIKTVIEYKYNNTSDERNRLE 517
 TKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQVDAIKTVIEYKYNNTSDERNRLE
 Sbjct: 61 TKWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQVDAIKTVIEYKYNNTSDERNRLE 120

Query: 518 SEYNINNIREELNKKVSLAMENIERFITESSIFYLMKLINEAKVSKLREYDEGVKEYLLD 577
 SEYNINNIREELNKKVSLAMENIERFITESSIFYLMKLINEAKVSKLREYDEGVKEYLLD
 Sbjct: 121 SEYNINNIREELNKKVSLAMENIERFITESSIFYLMKLINEAKVSKLREYDEGVKEYLLD 180

Query: 578 YISEHRSILGNSVQELNDLVTSTLNN SIPFELSSYTNDKILILYF 622
 YISEHRSILGNSVQELNDLVTSTLNN SIPFELSSYTNDKILILYF
 Sbjct: 181 YISEHRSILGNSVQELNDLVTSTLNN SIPFELSSYTNDKILILYF 225

>gi|134649|sp|P28760|SODM_BACCA Superoxide dismutase [Mn]
 gi|98098|pir|S22053 superoxide dismutase (EC 1.15.1.1) (Mn) - Bacillus caldotena
 gi|39453|emb|CAA44556.1| Manganese superoxide dismutase [Bacillus caldotenax]
 gi|7578509|gb|AAF64074.1|AF147780.1 superoxide dismutase [Geobacillus thermoleovo
 gi|384175|prf|1905285A superoxide dismutase
 Length = 204

Score = 387 bits (994), Expect = e-106
 Identities = 188/204 (92%), Positives = 188/204 (92%), Gaps = 16/204 (7%)

Query: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGH PDLQN----- 51
 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGH PDLQN
 Sbjct: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGH PDLQNKSLEELLSN 60

Query: 52 -----IRTAVRNNGGGHANHSLFWTILSPNGGGPEPTGELAEAINKKFGSFTAFKDEFS 104
 IRTAVRNNGGGHANHSLFWTILSPNGGGPEPTGELAEAINKKFGSFTAFKDEFS
 Sbjct: 61 LEALPESIRTAVRNNGGGHANHSLFWTILSPNGGGPEPTGELAEAINKKFGSFTAFKDEFS 120

Query: 105 KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNR RPE 164
 KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNR RPE
 Sbjct: 121 KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNR RPE 180

Query: 165 YIAAFWNIVNWDEVAKRYSEAKAK 188
 YIAAFWNIVNWDEVAKRYSEAKAK
 Sbjct: 181 YIAAFWNIVNWDEVAKRYSEAKAK 204

>gi|134650|sp|P00449|SODM_BACST Superoxide dismutase [Mn]
 gi|1070456|pir|DSBSNF superoxide dismutase (EC 1.15.1.1) (Mn) - Bacillus
 stearothermophilus
 gi|143203|gb|AAA22600.1| manganese superoxide dismutase (EC 1.15.1.1)
 gi|143552|gb|AAA22765.1| Mn-superoxide dismutase
 gi|143555|gb|AAA22767.1| Mn-superoxide dismutase
 Length = 204

Score = 386 bits (991), Expect = e-106

Identities = 186/204 (91%), Positives = 188/204 (92%), Gaps = 16/204 (7%)

Query: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQN----- 51
 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQN
 Sbjct: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKSLEELLSN 60
 Query: 52 -----IRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFS 104
 IRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELA+AINKKFGSFTAFKDEFS
 Sbjct: 61 LEALPESIRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELADAINKKFGSFTAFKDEFS 120
 Query: 105 KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE 164
 KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE
 Sbjct: 121 KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE 180
 Query: 165 YIAAFWNIVNWDEVAKRYSEAKAK 188
 YIAAFWN+VNWDEVAKRYSEAKAK
 Sbjct: 181 YIAAFWNVNVWDEVAKRYSEAKAK 204

>gi|7798571|dbj|BAA95631.1| superoxide dismutase [Geobacillus thermoleovorans]
 Length = 204

Score = 382 bits (981), Expect = e-105

Identities = 186/204 (91%), Positives = 186/204 (91%), Gaps = 16/204 (7%)

Query: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQN----- 51
 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQN
 Sbjct: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKSLEELLSN 60
 Query: 52 -----IRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFS 104
 IRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFS
 Sbjct: 61 LEALPESIRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFS 120
 Query: 105 KAAAGRFGSGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE 164
 KAAAGRF GWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE
 Sbjct: 121 KAAAGRFSPGWAWLVVNNGELEITSTPNQDSPIMEGKTPILGLDVWEHAYYLKYQNRPE 180
 Query: 165 YIAAFWNIVNWDEVAKRYSEAKAK 188
 YIAAFWNIVNWDEVAKRYSEAKAK
 Sbjct: 181 YIAAFWNIVNWDEVAKRYSEAKAK 204

>gi|448837|prf||1918164A superoxide dismutase
 Length = 205

Score = 381 bits (979), Expect = e-104

Identities = 186/205 (90%), Positives = 188/205 (91%), Gaps = 17/205 (8%)

Query: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQN----- 51
 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQN
 Sbjct: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQNKSLEELLSN 60
 Query: 52 -----IRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELAEAINKKFGSFTAFKDEFS 104
 IRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELA+AINKKFGSFTAFKDEFS
 Sbjct: 61 LEALPESIRTAVRNNGGGHANHSLFWTILSPNGGGEPTGELADAINKKFGSFTAFKDEFS 120
 Query: 105 KAAAGRFGSGWAWLVVNNGELEITSTPNQDS-PIMEGKTPILGLDVWEHAYYLKYQNRPE 163
 KAAAGRFGSGWAWLVVNNGELEITSTPNQDS PIMEGKTPILGLDVWEHAYYLKYQNRPE

Sbjct: 121 KAAAGRFGSGAWLVVNNGELEITSTPNQDSPPIMEGKTPILGLDVWEHAYYLKYQNR RP 180

Query: 164 EYIAAFWNIVNWDEVAKRYSEAKAK 188
EYIAAFWN+VNWDEVAKRYSEAKAK

Sbjct: 181 EYIAAFWNVNVNWDEVAKRYSEAKAK 205

>gi|7438650|pir|S48109 neurotoxin type F - Clostridium botulinum (fragment)
gi|407789|emb|CAA50147.1| botulinum neurotoxin type F [Clostridium botulinum]
gi|407791|emb|CAA50151.1| botulinum neurotoxin type F [Clostridium botulinum]
Length = 369

Score = 330 bits (847), Expect = 3e-89

Identities = 179/224 (79%), Positives = 193/224 (86%)

Query: 399 QKENFKEAFELLGAGILLEFVPELLIPTILVFTIKSFIGSSEXXXXXXXXXXXSLMERET 458
+K NF+EAFELLG GILLEFVPEL IP ILVFTIKS+I S E SL+ERE

Sbjct: 2 EKGNFEEAFELLGVGILLEFVPELTIPVILVFTIKSYIDSYENKNKAIKAINNSLIEREA 61

Query: 459 KWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQVDAIKTVIEYKYNNTSDERNRLES 518
KWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQVDAIKT IEYKYNNTSDE+NRLES

Sbjct: 62 KWKEIYSWIVSNWLTRINTQFNKRKEQMYQALQNQVDAIKTAIEYKYNNTSDEKNRLES 121

Query: 519 EYNINNIREELNKKVSLAMENIERFITESSIFYLMKLINEAKVSKLREYDEGVKEYLLDY 578
EYNINNI EELNKKVSLAM+NIERF+TESSI YLMKLINEAKV KL++YD VK LL+Y

Sbjct: 122 EYNINNIEELNKKVSLAMKNIERFMTESISYLMKLINEAKVGKLLKYDNHVKSDLLNY 181

Query: 579 ISEHRSILGNSVQELNDLVTSTLNNSIPFELSSYTNDKILILYF 622
I +HRSILG EL+DLVTSTLN+SIPFELSSYTNDKILI+YF

Sbjct: 182 ILDHRSILGEQTNELSDLVTSTLNSSIPFELSSYTNDKILIIYF 225

>gi|15613972|ref|NP_242275.1| superoxide dismutase [Bacillus halodurans]
gi|25285807|pir|A83826 superoxide dismutase sodA [imported] - Bacillus haloduran
C-125)
gi|10174025|dbj|BAE05128.1| superoxide dismutase [Bacillus halodurans]
Length = 202

Score = 327 bits (838), Expect = 4e-88

Identities = 153/202 (75%), Positives = 170/202 (84%), Gaps = 16/202 (7%)

Query: 1 MPFELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDL----- 49
M FELP LPYP +ALEPHID+ TMNIHH KHHNTYVT LNAALEGH L

Sbjct: 1 MAFELPKLPYPANALEPHIDEATMNIHHGKHHNTYVTKLNAALEGHSALAEKSIEALVSD 60

Query: 50 -----QNIRTAVERNNGGGHANHSLFWTILSPNGGGEGPTGELAEAINKKFGSFTAFKDEFS 104
+NIRTAVERNNGGGHANH+LFW ILSPNGGG PTGELA+AIN +FGSF FK++F+

Sbjct: 61 LDAVPENIRTAVERNNGGGHANHTLFWQILSPNGGGAPTGELADAINAEFGSFDQFKEKFA 120

Query: 105 KAAAGRFGSGAWLVVNNGELEITSTPNQDSPPIMEGKTPILGLDVWEHAYYLKYQNR RPE 164
AAA RFGSGAWLVVN+G+LEITSTPNQD+P+MEGKTPILGLDVWEHAYYL YQNR RP+

Sbjct: 121 DAAANRFGSGAWLVVNDGKLEITSTPNQDTPLMEGKTPILGLDVWEHAYYLNYQNR RPD 180

Query: 165 YIAAFWNIVNWDEVAKRYSEAK 186
YI+AFWN+VNWDEVAKRY+EAK

Sbjct: 181 YISAFWNVNVNWDEVAKRYNEAK 202

>gi|30264347|ref|NP_846724.1| superoxide dismutase, Mn [Bacillus anthracis str.
gi|30259005|gb|AAP28210.1| superoxide dismutase, Mn [Bacillus anthracis str. Ames
Length = 203

Score = 318 bits (815), Expect = 2e-85

Identities = 148/199 (74%), Positives = 163/199 (81%), Gaps = 16/199 (8%)

Query: 4 ELPALPYPYDALEPHIDKETMNIHHTKHHNTYVTNLNAALEGHPDLQN----- 51
ELP LPY YDALEPH DKETMNIHHTKHHNTY+TNLNAALEGH +L +

Sbjct: 5 ELPNLPYAYDALEPHFDKETMNIHHTKHHNTYITNLNAALEGHAELADKSVEELVANLNE 64

Animal Cell Lines

ATCC No.	Species	Name	Tissue Source	Cell Type	Appearance
CRL-11179	mouse	CATH.a	brain	neuron	
CCL-131	mouse	Neuro-2a	brain; neuroblastoma	neuroblast	neuronal and amoeboid
CCL-147	mouse	NB41A3	brain; neuroblastoma	neuroblast	neuroblast
CRL-2263	mouse	N1E-115	brain; neuroblastoma	neuroblast	fibroblast
HB-12317	mouse (neuroblastoma); rat (glioma)	NG108-15	brain; glioblastoma; neuroblastoma; hybrid	glial cell; neuron	flat; round
CRL-2534	mouse	C8-D30	brain; cerebellum	astrocyte, type III phenotype	neuronal
CRL-2535	mouse	C8-S	brain; cerebellum	astrocyte, type II phenotype	neuronal
CRL-2540	mouse	C8-B4	brain; cerebellum	microglial	neuronal
CRL-2541	mouse	C8-D1A	brain; cerebellum	astrocyte, type I phenotype	neuronal
CRL-1443	mouse	BC3H1	brain; smooth muscle like; tumor		
CRL-2299	mouse	bEnd.3	brain; cerebral cortex; endothelioma	endothelial	endothelial
CRL-2467	mouse	EOC 2	brain	microglia; macrophage	macrophage
CRL-2468	mouse	EOC 13.31	brain	microglia; macrophage	macrophage
CRL-2469	mouse	EOC 20	brain	microglia; macrophage	macrophage
CCL-107	rat	C6	brain; glioma	glial cell	fibroblast
CRL-2199	rat	C6/LacZ	brain; glioma	glial cell	fibroblast
CRL-2303	rat	C6/lacZ7	brain; glioma	glial cell	fibroblast
CRL-2397	rat	F98	brain; undifferentiated malignant glioma		glial
CRL-2433	rat	RG2 (D74)	brain; undifferentiated malignant glioma		glial
CRL-2200	rat	9L/lacZ	brain; gliosarcoma	glial cell	fibroblast
CRL-2005	rat	DI TNC1	brain; diencephalon; normal	astrocyte, type I phenotype	fibroblast
CRL-2006	rat	CTX TNA2	brain; cortex; normal	astrocyte, type I phenotype	fibroblast
CRL-1721	rat	PC-12	pheochromocytoma; adrenal gland	polygonal	
CRL-2032	cat	PG-4 (S+L-)	brain; normal	astrocyte	glial, astrocyte
CRL-2033	cat	G355-5	brain; normal	astrocyte	glial, astrocyte
CRL-1656	ferret	Mpf	brain; normal		
CRL-1700	ovine	SCP	brain; choroid plexus		
CRL-6538	ovine	OA1	brain; normal		fibroblast
CRL-2532	Japanese quail	QNR/D	neuroretina	neuronal	
CRL-2533	Japanese quail	QNR/K2	neuroretina	astroglial	

For more information or to place an order:

Phone: 800-638-6597

703-365-2700

Fax: 703-365-2750

E-mail: sales@atcc.org

Use our online shopping cart at:

www.atcc.org

These products are intended for laboratory research purposes only. They are not intended for use in humans.

© Copyright 2001. American Type Culture Collection. All rights reserved. ATCC® is a registered trademark of American Type Culture Collection.

ATCC No.	Species	Name	Tissue Source	Cell Type	Appearance
Non-tumor, neuronal-like cells					
CRL-10442	human	HCN-1A	brain	cortical neuron	neuronal
CRL-10742	human	HCN-2	brain	cortical neuron	neuronal
Tumor-derived neuronal-like cells					
CCL-127	human	IMR-32	brain; neuroblastoma	neuroblast	fibroblast; neuroblast
CRL-1718	human	CCF-STTG1	brain; astrocytoma		astrocytic
CRL-2060	human	PFSK-1	brain; cerebellum; malignant primitive neuroectodermal tumor	fibroblast	
CRL-2137	human	SK-N-AS	brain; neuroblastoma	neuroblast	epithelial
CRL-2142	human	SK-N-FI	brain; neuroblastoma	neuroblast	epithelial
CRL-2149	human	SK-N-DZ	brain; neuroblastoma	neuroblast	epithelial
CRL-2266	human	SH-SY5Y	brain; neuroblastoma		epithelial
CRL-2267	human	BE(2)-M17	brain; neuroblastoma		neuroblast
CRL-2268	human	BE(2)-C	brain; neuroblastoma		fibroblast
CRL-2270	human	MC-IXC	brain; neuroblastoma		neuroblast
CRL-2271	human	SK-N-BE(2)	brain; neuroblastoma		neuroblast
CRL-2273	human	CHP-212	brain; neuroblastoma		neuroblast
CRL-8621	human	SVG.p12	brain	astroglia	fibroblast
HTB-10	human	SK-N-MC	brain; neuroepithelioma, metastatic site: supra-orbital area		epithelial
HTB-11	human	SK-N-SH	brain; neuroblastoma, metastatic site: bone marrow		epithelial
HTB-12	human	SW 1088	brain; astrocytoma		fibroblast
HTB-13	human	SW 1783	brain; astrocytoma		fibroblast
HTB-15	human	U-118 MG	brain; glioblastoma; astrocytoma		mixed
Glioblastoma and glioma cells					
CRL-1620	human	A172	brain; glioblastoma		fibroblast
CRL-1690	human	T98G	brain; glioblastoma multiforme		fibroblast
CRL-2020	human	DBTRG-05MG	brain; glioblastoma	glial cell	fibroblast
CRL-2365	human	M059K	brain; malignant glioblastoma; glioma	glial cell	fibroblast
CRL-2366	human	M059J	brain; malignant glioblastoma; glioma	glial cell	fibroblast
CRL-7773	human	TE 615.T	brain; ganglioblastoma		epithelial
HTB-138	human	Hs 683	brain; glioma		fibroblast
HTB-14	human	U-87 MG	brain; glioblastoma; astrocytoma		epithelial
HTB-148	human	H4	brain; neuroglioma		epithelial
HTB-16	human	U-138 MG	brain; glioblastoma		polygonal
Medulloblastoma-derived cells					
CRL-8805	human	TE671 subline No. 2	brain; cerebellum; medulloblastoma		epithelial
HTB-185	human	D283 Med	brain; cerebellum; medulloblastoma, metastatic site: peritoneum		epithelial
HTB-186	human	Daoy	brain; cerebellum; desmoplastic cerebellar medulloblastoma		polygonal
HTB-187	human	D341 Med	brain; cerebellum; medulloblastoma		spheroid
Retinoblastoma-derived cells					
HTB-169	human	WERI-Rb-1	retinoblastoma; eye; retina		grape-like clusters of round cells
HTB-18	human	Y79	retinoblastoma; eye; retina		multicellular clusters
Neuroendocrine tissue					
CRL-5813	human	NCI-H660	lung; carcinoma; small cell lung cancer extrapulmonary origin (prostate), metastatic site: lymph node	neuroendocrine	epithelial
CRL-5893	human	NCI-H1770	lung; carcinoma; non-small cell lung cancer; metastatic site: lymph node	neuroendocrine	floating aggregates of round cells
CRL-2139	human	SK-PN-DW	malignant primitive neuroectodermal tumor; retroperitoneal embryonal tumor		epithelial
CRL-1973	human	NTERA-2 cl.D1	malignant pluripotent embryonal carcinoma; testis, metastatic site: lung		fibroblast